Investigating Advanced Magnetic Resonance Imaging for Improved Diagnosis and Prediction of Treatment Response in Wilms' Tumour Patients

Harriet J. Rogers

A thesis submitted to University College London
for the degree of

Doctor of Philosophy in Imaging and Biophysics

Declaration

I, Harriet Rogers, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

Signed:

Abstract

Wilms' tumour is the most common paediatric renal tumour. In Europe treatment involves pre-operative chemotherapy followed by surgery. Many patients receive MRI scans which include diffusion weighted imaging (DWI) throughout their diagnosis and treatment.

This thesis retrospectively acquired Wilms' tumour MRI data and prospectively acquired renal DWI data in healthy volunteers. Four models of diffusion were used throughout this thesis; mono-exponential, IVIM (intravoxel incoherent motion), stretched exponential, and kurtosis.

In healthy volunteers, models were compared based on the reproducibility of the parameters, when calculated based on different b values and magnetic fields strengths. It was shown that ADC, D (IVIM), f (IVIM), D_k (kurtosis), and α (stretched exponential) had high levels of reproducibility whereas reproducibility was poorer in D^* (IVIM), K (kurtosis) and DDC (stretched exponential).

Model fits were compared in Wilms' tumour and contralateral normal kidney data using the Akaike Information Criterion. It was shown that all raw DWI data favoured non-Gaussian models as opposed to a mono-exponential model. DWI data acquired in Wilms' tumour favoured the stretched exponential model, and DWI data acquired in normal kidneys favoured the IVIM model.

The volume of necrotic tissue post-chemotherapy is an important marker of treatment response. However, currently identification of necrosis relies on gadolinium contrast enhancement. It was shown that a combination of T_1weighted imaging and ADC could provide an alternative method to visualising and quantifying necrosis, allowing future studies to estimate the volume fraction of necrosis in Wilms' tumour without gadolinium.

Finally, it was shown that certain Wilms' tumour subtypes could be distinguished *in vivo* using DWI, whereas currently this relies on histological tissue analysis post-surgery. The parameters D^* (IVIM) and K (kurtosis) provided the best stratification between subtypes, however, the earlier study demonstrated that the reproducibility of these parameters was poor, which may limit their clinical utility.

Acknowledgements

There are many people who must be acknowledged, as without them, this thesis would not have been possible. Firstly, a massive thank you to my supervisors Chris Clark and Patrick Hales, who interviewed me and gave me the opportunity and privilege of undertaking this PhD. Both have been incredibly supportive and helpful throughout this project, offering advice and suggestions along the way, and also being supportive of my personal circumstances. Patrick has been there for me on a daily basis and so for all the MATLAB script fixings, paper revisions, meetings with clinicians, scanner sessions, trips around the country and world, coffee chats and support for when I went through personal difficulties, thank you.

I would also like to thank Kathy Pritchard-Jones for her advice on clinical questions in Wilms' tumour and to Jesper Brok who provided me with guidance and help in framing investigations. This leads me on to thank Susie Shelmerdine and Martijn Verhagen who worked incredibly hard on our paper, and it would not have been possible without their time, effort, and expertise.

Thank you to Ramneek, who identified patient lists and talked me through the PACs system. Thank you to all of the radiographers, who worked tirelessly acquiring the data and to Tina and Paul in particular, who helped with the scanning of all my volunteers, and offered advice on protocols. Thank you, also, to all the patients and families who allowed me to undertake this work.

It has been an amazing experience working in the developmental imaging and biophysics section, and I would not have got through the last three years without the support of my incredible colleagues. Thank you especially to Winok and Hanne, who were always there to pick me up when things got too difficult and always happy to grab a glass of pinot to celebrate our successes.

Thank you to Shumi, who volunteered for many months and spent hours drawing ROIs, it is a long repetitive task and I am so grateful for your assistance. Thank you not only for your help on the project but for being an incredible friend over the last eight years, for all the laugher and being there for the good and bad times. This leads me on to thank all of my friends who have listened to me talk about academia and my goals since we met in 2010: Ellie, Nikki and Charlie, a group of friends who will always be there for

you and will always bring a smile to your face, and who are more than happy to come see you in Paris when you are at a conference.

To my oldest friends Amy and Sarah, thank you for always listening to me and putting up with me for the highs and lows of my academic career. You always ask when I will stop being a student and here we are.

Thank you to my family who supported my decision to go into academia and have been there through everything, thank you for giving me a place to live, for providing emotional support and for offering advice on my projects. Without your love and guidance, I would not be where I am today, for everything, thank you.

Lastly thank you to Ben. You have been there for all the good and bad times, you have listened to me go on about the different projects, results and methodological issues, thank you for your support and giving me the best three years of my life, this is for you.

Impact Statement

This thesis investigated the use of diffusion weighted imaging in Wilms' tumours. A range of advanced non-Gaussian models of diffusion were compared. Each model provides various biomarkers of diffusion and have the potential to provide clinically useful information.

This thesis demonstrated that in Wilms' tumour tissue and normal renal tissue, diffusion does not exhibit mono-exponential behaviour, but is better described by non-Gaussian models. In Wilms' tumour the stretched exponential model provided the best description of the data and in normal renal tissue the IVIM model (intravoxel incoherent motion) provided the best description. This finding could be beneficial in many areas of research as many studies utilise ADC (a diffusion parameter derived from a mono-exponential model) and thus assume the tissue has a mono-exponential signal decay as the diffusion weighting is increased. Being aware that tissue may be better described by other models may lead to more accurate representations of the underlying tissue structure.

Another finding was that many parameters produced from diffusion models are reproducible, including ADC, D (IVIM), f (IVIM), D_k (kurtosis), and α (stretched exponential), when generated based on acquisitions using different b values and at different magnetic field strengths. This finding will help with generalising the results from this thesis and interpreting research from different centres who do not have identical imaging protocols. Furthermore, some parameters were less reproducible [D^* (IVIM), K (kurtosis) and DDC (stretched exponential)], thus aiding the interpretation of future research using these parameters, as some findings may not be able to be replicated across different centres.

A key result from this thesis was the ability to identify necrotic tissue within Wilms' tumour using a non-invasive method. It was shown that necrotic tissue could be visualised and quantified using a combination of ADC and T_1weighted imaging, whereas currently an injection of gadolinium contrast is required. This finding has the potential to be clinically useful as the volume of necrotic tissue post-chemotherapy is informative of treatment response, and gadolinium is not always appropriate in some patients.

Furthermore, this thesis highlighted that subtypes of Wilms' tumour may be able to be identified *in vivo* using diffusion measurements, whereas currently this can only be determined via histological analysis after surgery. This thesis suggested that non-Gaussian

diffusion parameters may be superior to ADC in distinguishing different Wilms' tumour subtypes.

Therefore, overall this thesis has suggested that non-Gaussian models of diffusion should be investigated in future research in Wilms' tumour, and potentially other forms of abdominal cancer. These models provided an improved fit to the raw DWI signal and provided novel biomarkers to describe the microstructure and physiology of the tumour tissue.

Papers and Abstracts Arising from This Thesis

Publications:

2018: Rogers HJ, Verhagen MV, Shelmerdine SC, Clark CA, & Hales PW. (2018) An alternative approach to contrast-enhanced imaging: diffusion-weighted imaging and T_1-weighted imaging identifies and quantifies necrosis in Wilms tumour. *European Radiology,* doi: 10.1007/s00330-018-5907-z.

Abstracts Accepted for Oral Presentations:

2016: Rogers HJ, Hales PW, Olsen Ø, Sebire N, Pritchard-Jones K, & Clark CA. (2016). Is a Mono-Exponential Fit Suitable for Diffusion Data in Wilms Tumour? A Multi-Model Comparison Study. *BC-ISMRM Annual Scientific Meeting, Leeds*

2017: Rogers HJ, Hales PW, Yoheswaran S, Verhagen MV, & Clark CA. (2017). Identifying Necrosis in Wilms Tumours using Diffusion MRI. *BC-ISMRM Annual Scientific Meeting, Liverpool*

Abstracts Accepted for Power Pitch:

2017: Rogers HJ, Hales PW, Pritchard-Jones K, & Clark CA. (2017). Diffusion-Weighted MRI Identifies Viable Tissue in Wilms Tumour: Application for Subtype Analysis and Response to Chemotherapy. *ISMRM Annual Scientific Meeting, Hawaii*

Abstracts Accepted for Electronic Poster:

2018: Rogers HJ, Verhagen MV, Shelmerdine SC, Clark CA, & Hales PW. (2018). ADC and T_1-weighted MRI Identifies Necrosis in Wilms Tumours Without the Need for Contrast Agents. *ISMRM Annual Scientific Meeting, Paris*

Table of Contents

Figure List

Table List

Chapter 1: Wilms' Tumour

This chapter discusses Wilms' tumour, highlighting the aetiology and histological presentation of the disease. Additionally, this chapter explains the differences in treatment approaches used across the world and recent updates in protocols. Throughout this thesis patients with Wilms' tumour are referred to as Wilms' tumour patients.

1.1: Wilms' Tumour in Context

Wilms' tumour, otherwise known as nephroblastoma, accounts for approximately 5% of all cancers in children under 15 years of age[1]. Wilms' tumour is the most common paediatric renal tumour equating to approximately 80-90% of all kidney tumours[1]. The majority of these tumours are unilateral; affecting only one of the kidneys, but around 7% of all cases are bilateral[1]. The majority of diagnoses are in those between 3 and 4 years old, with a decline in frequency being seen in children over 5 years; however Wilms' can present in older children and even in adults in very rare cases[1].

1.2: Wilms' Tumour Aetiology

The kidneys are formed through the development of nephrons from foetal metanephric blastema which surrounds the ureteric bud[2]. In healthy development the foetal renal tissue will mature during gestation, however in some cases foetal tissue can remain into infancy in the form of nephrogenic rests[2]; these rests represent clusters of abnormally persistent immature embryonal cells[3]. Nephrogenic rests can be sub classified as either perilobar nephrogenic rests or intralobar nephrogenic rests; the former being limited to the renal cortex, and the latter occurring throughout the renal lobe[3]. Nephrogenic rests are thought to be precursors to Wilms' tumour[4]; in that these rests can undergo transformation into malignant renal tumour[2], although only approximately 1% of children with nephrogenic rests go on to develop Wilms' tumours[5], other rests will become hyperplastic, dormant or spontaneously involute[6]. Nephrogenic rests can often be confused with Wilms' tumour, particularly in the cases of bilateral tumours; as nephrogenic rests have been found in 94-99% of those with bilateral Wilms' tumour[3].

In a small number of Wilms' tumour patients all or part of the WT-1 gene is deleted, and it was suggested that metanephric stem cells cannot differentiate to normal

kidney components when WT-1 is inactivated[7]. However, the majority of Wilms' tumour patients do not have this gene mutation and the genetic components of Wilms' tumour is much more complex[8]. Around 1-2% of cases are familial[9], but in most cases the causes of Wilms' tumour are unknown.

1.3: Clinical and Histological Presentation

Wilms' tumour presents clinically as a large painless abdominal mass, with only around a third of patients exhibiting abdominal pain, vomiting, and tiredness[6]. Studies have shown that around 25% of patients have hypertension, and other abnormalities and syndromes such as genitourinary malformations, aniridia, Beckwith-Wiedemann syndrome, and Denys-Drash syndrome, are also present in around 13-28% of patients[6].

Histologically, Wilms' tumours are comprised of blastemal, stromal and epithelial cells; the microscopic appearance will differ depending on the amount of each of these components. In addition to these components Wilms' tumours can be characterised as having 'favourable' or 'unfavourable' histology; this is related to the absence or presence of anaplasia (cells with poor differentiation, orientation, and abnormal morphological characteristics compared to mature cells). Anaplastic Wilms', or unfavourable Wilms', accounts for around 5% of cases, and 5 year event-free survival rates are much lower than those with a favourable histology, with rates of around 40-50%[10,11], compared to approximately 90% in those with favourable Wilms'[12]. However, the amount of anaplasia is correlated with patient outcome, and those with minimal localised (focal) anaplasia have similar outcomes to favourable Wilms' patients, whereas as those with diffuse anaplasia have far worse outcomes[2]. Anaplastic Wilms' are more resistant to chemotherapy, however these tumours are not aggressive in terms of tumour growth[10]. Additionally, it is the older patients who are more likely to have anaplasia and thus if older children with nephrogenic rests develop a mass, it has been suggested that immediate surgical resection should be performed[2].

1.4: Approaches to Treatment

There are two approaches used to characterise and investigate Wilms' tumour, these are the Children's Oncology Group (COG) - formally known as the National Wilms' Tumor Study Group (NWTSG), and the Société Internationale d'Oncologie Pédiatrique

(SIOP) – the International Society for Paediatric Oncology. It is the combined efforts of these approaches' which have led to survival rates of approximately 90%[12].

The COG approach favours nephrectomy, where possible, followed by post-operative chemotherapy. This approach provides access for immediate histological confirmation and analysis of tumour tissue; providing information on the presence and proportion of cell composition. The SIOP approach, however, advocates pre-operative chemotherapy, aiming to reduce the size of the tumour; resulting in easier removal with fewer tumour ruptures during surgery. The SIOP approach is used throughout Europe, and the following details on staging and classification are based on this approach.

1.5: Wilms' Tumour Stages and Subtypes

Wilms' tumours can be categorised by their stage which is dependent on their extension and surgical outcomes; details of these stages as described by SIOP[13] are given Table 1.1.

Table 1.1. Wilms' tumour stages as described by the SIOP-2001 protocol

Stage	Staging as Described by SIOP-2001
Stage I	The tumour affects one kidney, it has not spread and was completely removed during surgery
Stage II	The tumour has spread beyond the kidney, but no cancerous cells are found in more distant organs such as the lungs. The tumour was completely removed in surgery
Stage III	The tumour was not completely removed in surgery, and/or it has spread to lymph nodes in the abdomen, and/or it ruptured prior to or during surgery
Stage IV	The tumour has spread to distant organs beyond the abdomen, commonly this is the lungs, however it can also include the liver, bone or brain
Stage V	The tumour is bilateral and affects both kidneys

In addition to staging, the revised SIOP-2001 protocol also classified Wilms' tumours as low risk, intermediate risk and high risk[13], based on their histological appearance. Table 1.2 provides details of this classification system.

Table 1.2. Wilms' tumour risk type classification as described by SIOP-2001

Risk Group	Description
Low Risk	Cystic partially differentiated nephroblastoma
	Completely necrotic
Intermediate Risk	Epithelial Subtype
	Stromal Subtype
	Mixed Subtype
	Regressive Subtype
	Focal Anaplasia
High Risk	Blastemal Subtype
	Diffuse Anaplasia

Treatment is based on stages and risk groups; for those with stage I the treatment is 4 weeks of chemotherapy followed by surgery to remove the kidney, if the tumour is low risk then this is the end of treatment, however, if the tumour is intermediate or high risk then more chemotherapy may be necessary following surgery. For stage II low and intermediate risk tumours, chemotherapy is usually given for 4 weeks followed by surgery after which the patient may undergo further chemotherapy for 4 or 27 weeks depending on the risk category of the tumour. If the tumour is stage II with a high-risk type, the treatment is the same, except post-operative chemotherapy is more intense; for 34 weeks. Stage III patients will have a similar treatment plan to stage II patients, however, the post-operative chemotherapy will be dependent on the risk type and the patient may also receive post-operative radiotherapy. Stage IV patients will undergo 6 weeks of pre-operative chemotherapy, and the treatment following surgery will be dependent on the risk type and how well the patient is responding to the previous treatment. The treatment plan for stage V is pre-operative chemotherapy, followed by surgery to remove as much of

the tumour as possible from each kidney. Post-operative chemotherapy and radiotherapy will then be dependent on surgical outcome, risk category and treatment response.

In the UK a percutaneous biopsy is performed prior to chemotherapy to confirm the Wilms' diagnosis (this does not occur in other SIOP countries)[14]. After chemotherapy and surgery, a more accurate diagnosis is given, and the tumour will be assigned to a histological subtype and risk group, as well as a stage.

A subtype is given to a Wilms' tumour based on histological features. Firstly, the percent of necrosis and regressive change as a result of chemotherapy is determined, and if only necrotic tissue is present, and no viable tumour remains, then the tumour is labelled as completely necrotic. If more than two thirds of the tumour has undergone regressive change then it will be labelled as regressive. If less than two thirds of the entire mass has undergone regressive change then the subtype is classified based on the predominant viable cell type; either stromal, epithelial or blastemal. If a tumour has a range of cells, none of which having a 66% majority then it is classed as mixed.

Table 1.3. Wilms' tumour subtypes as described by SIOP-2001

Tumour Subtype	**Histological Features of Tumour**			
	Chemotherapy Induced Changes	Epithelial Tissue	Stromal Tissue	Blastemal Tissue
Completely Necrotic	100%	0%	0%	0%
Regressive	>66%	0-33%	0-33%	0-33%
Mixed	<66%	0-65%	0-65%	0-65%
Epithelial	<66%	66-100%	0-33%	0-10%
Stromal	<66%	0-33%	66-100%	0-10%
Blastemal	<66%	0-33%	0-33%	66-100%

It is important to be aware of the distinction between a stromal or epithelial subtype, as based on the SIOP-2001 classification, and a stromal or epithelial predominant tumour, as based on the COG classification. For example, the SIOP approach classifies a tumour as stromal if stromal cells make up more than 66% of the viable tumour but the remaining viable tumour must have less than 10% blastema, otherwise it is classed as a mixed type. In the COG approach, the remaining viable tumour can have up to 33% blastemal and still be regarded as a stromal predominant tumour. Table 1.3 provides further details of how the SIOP-2001 categorises Wilms' tumour subtypes[13].

There are some issues with the classification system which were highlighted by Vujanić and Sandstedt[15]. For example, defining necrotic tissue is difficult as there are no clear criteria for defining tissue as having no viability. The authors highlighted that it can be easy to recognise 'dead' cells, but it is much more difficult to define the cells which are 'dying'. Necrotic tissue can be visualised using MRI (magnetic resonance imaging) and a gadolinium-based contrast agent; where areas of reduced signal represent necrotic tissue. However, gadolinium has many associated side effects such as vomiting, headaches and irritation. It also requires venous access; thus, it may not always be appropriate. Additionally, gadolinium can accumulate in neuronal tissue, however the resulting pathology of this is unknown[16]. Further details of studies which have used gadolinium contrast-enhanced MRI in Wilms' tumour can be found in Chapter 4 and an investigation into an alternative approach for identifying necrosis in Wilms' tumour can be seen in Chapter 8.

Furthermore, the most difficult problem associated with the SIOP-2001 classification system is determining whether a tumour has over 66% of a certain cell type. This can have implications for treatment, for example, a mixed type Wilms' tumour with lots of blastemal tissue is part of the intermediate risk category, however a blastemal type is a high-risk tumour, and the treatment for these two types differs. Currently classification is semi-quantitative, based on looking at a few tumour slides and generating estimates. However, Wilms' tumours can be highly heterogeneous and cellular composition can differ per slice, thus based on the current methods it is possible that errors could be made when classifying the tumour subtype. It would thus be very beneficial to able to non-invasively determine the volume of necrotic tissue and assess the histological subtype over the entire the tumour.

When considering the subtypes of Wilms' tumour is it important to keep in mind that the different approaches (either pre-operative chemotherapy or immediate surgery) may also alter the presence of certain subtypes in patient populations. For example, in

early SIOP trials in which immediate surgery was performed the most common subtypes were mixed followed by blastemal and epithelial predominant types. However, when pre-operative chemotherapy was used the most common type was regressive, then mixed, stromal, blastemal and epithelial predominant, and also a small proportion of completely necrotic tumours. It has been suggested that that the chemotherapy destroys the blastemal and epithelial components, and induces maturation of the existing stromal tissue, thus more stromal subtypes are likely to be seen in pre-operative chemotherapy patient groups[15].

The blastemal subtype has been a point of discussion over the years, with reports suggesting that it is one of the most responsive subtypes to chemotherapy[17], as this subtype has shown to have a greater decrease in volume following treatment compared to other subtypes[18]. However, the presence of blastema after chemotherapy indicates a poorer outcome[19]. It has been suggested that rather than calculating the percent of blastemal cells which remain after treatment, instead the absolute volume of blastemal tissue should be considered, with a threshold of 20–50 ml of remaining blastemal tissue being classed as high-risk[20].

1.6: Survival Rates

Survival rates in patients with stromal, epithelial, mixed and regressive subtypes are generally high, with approximately 89% of patients having relapse-free survival at 4 year follow-ups[19]. However, as stated earlier, the blastemal subtype has a worse outcome; with patients experiencing approximately 62% relapse-free survival at 5 year follow-ups[21]. As the blastemal subtype is higher risk than epithelial, stromal, mixed, and regressive, it was suggested that treatment should be intensified for blastemal subtypes and perhaps a milder treatment plan used for the intermediate subtypes[19]. To investigate this, survival rates of unilateral blastemal subtypes from SIOP-2001, where an intensified post-operative chemotherapy was used, were compared to the survival rates from unilateral blastemal subtypes from previous SIOP protocols (1993-2001)[18]. It was shown that 5-year event free survival rates increased from 67% to 80% when the more intense therapy was used. However, overall survival rates did not significantly differ, increasing only from 84% to 88% across all blastemal subtypes. Although, when analysis was constrained to stage I blastemal subtypes overall survival did increase from 71% to 100%, therefore survival is heavily dependent on both staging and histological subtype.

While overall survival rates and event free survival rates in Wilms' tumour are relatively high, it is important to note the longer-term effects of treatment. It was reported that 65.4% of Wilms' tumour patients developed later chronic health problems following chemotherapy when assessed at a 25 year follow-up, with 24.2% reporting severe health conditions[22]. Thus, adapting treatment based on subtypes, risk and staging, and potentially reducing treatment in lower risk groups may alleviate future health problems.

1.7: Recent Updates: The UMBRELLA protocol

The UMBRELLA-2016 protocol was recently introduced to replace the SIOP-2001 protocol. It aims to incorporate data from histology and imaging to implement a more standardised treatment, and also focus on genetic factors[23]. In addition to Wilms' tumours, non-Wilms' renal tumours are also addressed. However, diagnosis and treatment of Wilms' tumour subtypes have not significantly changed since SIOP-2001. Although, currently still under review, it is likely that the blastemal subtype will be characterised using absolute volume of blastema rather than percentage as was previously suggested[20]. Additionally, this higher-risk subtype will continue to receive more intensive post-operative chemotherapy, based on the results from SIOP-2001[18].

The UMBRELLA protocol is also similar to SIOP-2001 in its recommendations for children under 6 months; suggesting that they have immediate surgery instead of pre-operative chemotherapy, and that histological assessment prior to treatment is not advised as it delays treatment[23].

Stage IV Wilms' tumours, in which the tumour has spread to other locations, was found in approximately 17% of patients[23], and the UMBRELLA protocol recommends 6 weeks of pre-operative treatment, similar to the SIOP-2001 approach, as it was found that 61-67% of patients had a complete metastatic response[24,25]. Additionally, the UMBRELLA protocol suggests that individual centres discuss the best treatment path for those with stage IV Wilms'.

For bilateral Wilms', stage V, the UMBRELLA protocol suggests avoiding total nephrectomy as it has been shown that bilateral nephron sparing surgery provided better functional renal outcomes compared to other surgeries[26], this once again is similar to the SIOP-2001 approach. However, unlike the SIOP-2001 approach, the UMBRELLA protocol suggests that pre-operative chemotherapy in bilateral Wilms' is limited to a maximum of 12 weeks, with evaluation of treatment at 6 weeks[23]. Evaluation is likely to be based on

imaging techniques such as MRI; which is a non-invasive tool that has the potential to assess tissue status.

The UMBRELLA protocol also details guidelines for diagnoses of Wilms' tumour in adults[23]. As adult Wilms' tumour is incredibly rare, and diagnosis is frequently made unexpectedly after nephrectomy, there are often substantial delays in treatment. If histological diagnosis can be made prior to surgery then pre-operative chemotherapy is also recommended, however these cases are exceptionally rare.

Summary of Wilms' Tumour

Wilms' tumour is the most common form of renal tumour in childhood, and in general, survival rates are relatively high, but this is dependent on histology, staging, and risk group. Subtypes of Wilms' can be classified as low, intermediate or high risk based on their histological features. Treatment in Europe involves pre-operative chemotherapy to reduce tumour size, followed by surgery. The protocol is known as SIOP, however very recently the name has been changed along with a few differences in treatment to UMBRELLA.

The current challenges seem to be the identification of necrotic tissue and subtypes. MRI may be able to aid in these investigations as it is a non-invasive tool which allows one to visualise the entire tumour and inspect it on a slice-by-slice basis. Additionally, MRI can provide quantitative information potentially indicating the cellular density of tissues. By analysing tumour tissue using MRI, necrosis and individual subtypes may be able to be identified which in turn will impact treatment decisions and prognosis as described in Sections 1.5 and 1.6. Furthermore, the findings may be able to be applied to other abdominal tumours such as neuroblastoma.

This thesis focused on the use of MRI in Wilms' tumours, and data was collected from Great Ormond Street Hospital, London. It is one of the UK's most prominent Wilms' tumour referral centres and sees around 25% of all Wilms' tumour patients, making it an excellent centre for these investigations.

The following chapters describe MRI and various models of diffusion weighted imaging (DWI) and look into how these techniques have been previously used in Wilms' tumour. The later chapters include original research which investigate these models in healthy kidney tissue and in Wilms' tumours.

References: Chapter 1

1. Pastore, G. *et al.* Malignant renal tumours incidence and survival in European children (1978-1997): report from the Automated Childhood Cancer Information System project. *Eur. J. Cancer Oxf. Engl. 1990* **42**, 2103–2114 (2006).
2. Lonergan, G. J., Martínez-León, M. I., Agrons, G. A., Montemarano, H. & Suarez, E. S. Nephrogenic rests, nephroblastomatosis, and associated lesions of the kidney. *RadioGraphics* **18**, 947–968 (1998).
3. Beckwith, J. B., Kiviat, N. B. & Bonadio, J. F. Nephrogenic rests, nephroblastomatosis, and the pathogenesis of Wilms' tumor. *Pediatr. Pathol.* **10**, 1–36 (1990).
4. Beckwith, J. B. Precursor lesions of Wilms tumor: clinical and biological implications. *Med. Pediatr. Oncol.* **21**, 158–168 (1993).
5. Breslow, N. *et al.* Ethnic variation in the incidence, diagnosis, prognosis, and follow-up of children with Wilms' tumor. *J. Natl. Cancer Inst.* **86**, 49–51 (1994).
6. Kalapurakal, J. A. *et al.* Management of Wilms' tumour: current practice and future goals. *Lancet Oncol.* **5**, 37–46 (2004).
7. Re, G. G., Hazen-Martin, D. J., Sens, D. A. & Garvin, A. J. Nephroblastoma (Wilms' tumor): a model system of aberrant renal development. *Semin. Diagn. Pathol.* **11**, 126–135 (1994).
8. Ruteshouser, E. C., Robinson, S. M. & Huff, V. Wilms tumor genetics: Mutations in WT1, WTX, and CTNNB1 account for only about one-third of tumors. *Genes. Chromosomes Cancer* **47**, 461–470 (2008).
9. Matsunaga, E. Genetics of Wilms' tumor. *Hum. Genet.* **57**, 231–246 (1981).
10. Beckwith, J. B., Zuppan, C. E., Browning, N. G., Moksness, J. & Breslow, N. E. Histological analysis of aggressiveness and responsiveness in Wilms' tumor. *Med. Pediatr. Oncol.* **27**, 422–428 (1996).
11. Green, D. M. *et al.* Treatment of children with stages II to IV anaplastic Wilms' tumor: a report from the National Wilms' Tumor Study Group. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **12**, 2126–2131 (1994).
12. Dome, J. S., Perlman, E. J. & Graf, N. Risk stratification for wilms tumor: current approach and future directions. *Am. Soc. Clin. Oncol. Educ. Book Am. Soc. Clin. Oncol. Annu. Meet.* 215–223 (2014). doi:10.14694/EdBook_AM.2014.34.215

13. Vujanić, G. M. *et al.* Revised International Society of Paediatric Oncology (SIOP) working classification of renal tumors of childhood. *Med. Pediatr. Oncol.* **38**, 79–82 (2002).
14. Vujanić, G. M., Kelsey, A., Mitchell, C., Shannon, R. S. & Gornall, P. The role of biopsy in the diagnosis of renal tumors of childhood: Results of the UKCCSG Wilms tumor study 3. *Med. Pediatr. Oncol.* **40**, 18–22 (2003).
15. Vujanić, G. M. & Sandstedt, B. The pathology of Wilms' tumour (nephroblastoma): the International Society of Paediatric Oncology approach. *J. Clin. Pathol.* **63**, 102–109 (2010).
16. McDonald, R. J. *et al.* Intracranial Gadolinium Deposition after Contrast-enhanced MR Imaging. *Radiology* **275**, 772–782 (2015).
17. Zuppan, C. W., Weeks, D. A., Beckwith, J. B., Luckey, D. W. & Pringle, K. C. The effect of preoperative therapy on the histologic features of wilms' tumor. An analysis of cases from the third national Wilms' tumor study. *Cancer* **68**, 385–394 (1991).
18. van den Heuvel-Eibrink, M. M. *et al.* Outcome of localised blastemal-type Wilms tumour patients treated according to intensified treatment in the SIOP WT 2001 protocol, a report of the SIOP Renal Tumour Study Group (SIOP-RTSG). *Eur. J. Cancer* **51**, 498–506 (2015).
19. Weirich, A. *et al.* Clinical impact of histologic subtypes in localized non-anaplastic nephroblastoma treated according to the trial and study SIOP-9/GPOH. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **12**, 311–319 (2001).
20. Graf, N. *et al.* Is the absolute blastema volume after preoperative chemotherapy in nephroblastoma relevant for prognosis? [Abstract]. *Pediatr. Blood Cancer* **57**, 741–742 (2011).
21. Graf, N. *et al.* Characteristics and outcome of stage II and III non-anaplastic Wilms' tumour treated according to the SIOP trial and study 93-01. *Eur. J. Cancer Oxf. Engl. 1990* **48**, 3240–3248 (2012).
22. Termuhlen, A. M. *et al.* Twenty-five year follow-up of childhood Wilms tumor: a report from the Childhood Cancer Survivor Study. *Pediatr. Blood Cancer* **57**, 1210–1216 (2011).
23. Heuvel-Eibrink, M. M. van den *et al.* Position paper: Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP–RTSG 2016 protocol. *Nat. Rev. Urol.* **14**, 743–752 (2017).
24. Verschuur, A. *et al.* Treatment of pulmonary metastases in children with stage IV nephroblastoma with risk-based use of pulmonary radiotherapy. *J. Clin. Oncol. Off. J. Am. Soc. Clin. Oncol.* **30**, 3533–3539 (2012).

25. Warmann, S. W. *et al.* Tumor biology influences the prognosis of nephroblastoma patients with primary pulmonary metastases: results from SIOP 93-01/GPOH and SIOP 2001/GPOH. *Ann. Surg.* **254**, 155–162 (2011).
26. Aronson, D. C., Slaar, A., Heinen, R. C., de Kraker, J. & Heij, H. A. Long-term outcome of bilateral Wilms tumors (BWT). *Pediatr. Blood Cancer* **56**, 1110–1113 (2011).

Chapter 2: Fundamentals of MRI

Magnetic Resonance Imaging (MRI) is a safe, non-invasive tool which allows us to study the inner workings of the human body. This chapter focuses on explaining the fundamentals and basic principles of MRI. Information in this chapter is based on several textbooks and articles, which are given in references 1 – 7.

2.1: Proton Spins and Magnetisation

The human body is primarily made up of water (over 60%), with the main component of water being hydrogen. It is this abundance of hydrogen in the human body which forms the basis of MRI. A hydrogen nucleus, or a single positively charged proton, rotates around its axis. As this is a moving charge, it has its own very small magnetic field, which is known as a magnetic moment. When a proton is placed in an external magnetic field (known as B_0) it attempts to align itself to the direction of this magnetic field. Due to the laws of quantum mechanics it cannot perfectly align and so continues to spin, or precess, around the direction of the field. The precessional frequency is proportional to the external magnetic field, the higher the external field the higher the frequency, and this can be given by the Larmor equation (Equation 2.1):

$$\omega = \gamma B_0 \quad [2.1]$$

ω = Precession/Larmor Frequency

γ = the gyromagnetic ratio (a constant = 42.57Mhz T^{-1} for hydrogen nuclei)

B_0 = the external magnetic field

Protons can be orientated in any direction, creating a near-spherical distribution although there is tendency to point along the external magnetic field. The application of this magnetic field also makes the proton distribution precess around this fixed axis. An example is shown in Figure 2.1.

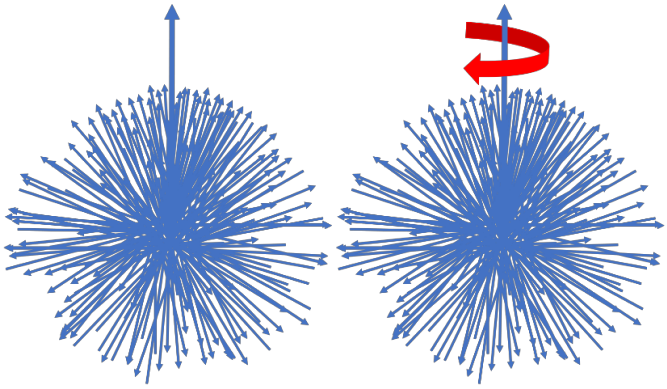

Figure 2.1. The distribution of spins is almost spherical, with tendency to point in the direction of the external field, shown here by more arrows pointing towards the larger arrow (the applied field). The field also makes this distribution precess about the direction of this field, shown with the red arrow.

Due to the vast number of protons in the human body, we can represent the spins as a vector of the average magnetic moments rather than look at each proton individually. The sum of all the spins is known as the net magnetization, M_0, and is aligned with the main magnetic field (B_0). For convention, it is shown in the z-direction. This magnetisation is very small compared to the main external magnetic field. When M_0 is parallel to B_0 it cannot be measured, and thus it is flipped 90° into the x-y plane, also known as the transverse plane, as shown in Figure 2.2. To flip M_0, a radio frequency pulse (RF pulse) is used, which is applied at the Larmor frequency.

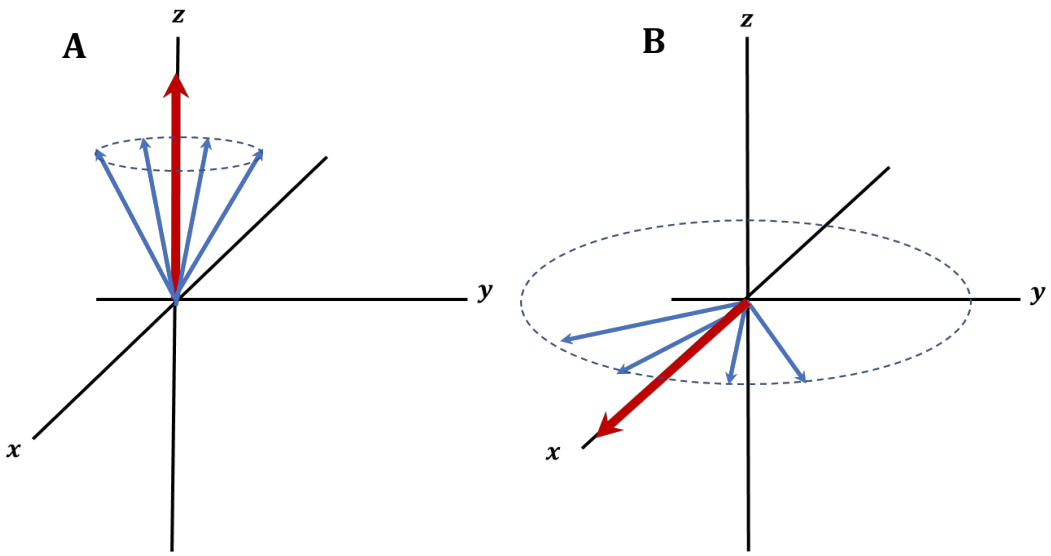

Figure 2.2. (A) The net magnetisation (red arrow) precessing around the z-axis, B_0. **(B)** The net magnetisation after the application of the radio frequency pulse so that it is in the transverse plane.

This M_0 magnetisation is now precessing in the transverse plane and can thus be measured. This precession induces a voltage which is detected by receiver coils (loops of wire) via electromagnetic induction. The coil detects the magnetic field which oscillates at the Larmor frequency. The induced signal decays exponentially as the protons start to dephase with respect to one another, which is termed the free induction decay (FID). However, the FID is not measured directly but rather echoes are created.

2.2: Echoes

There are two common types of echoes: gradient echo (GE) and spin echo (SE). For a SE sequence a 90° RF excitation pulse is applied, the spins thus flip into the transverse plane and begin to dephase. After a certain time, a 180° pulse is applied which then reverses the dephasing of the spins (albeit not completely), forming the spin echo, as shown in Figure 2.3. (The 'TE' in this diagram is the echo time, which is explained in section 2.6).

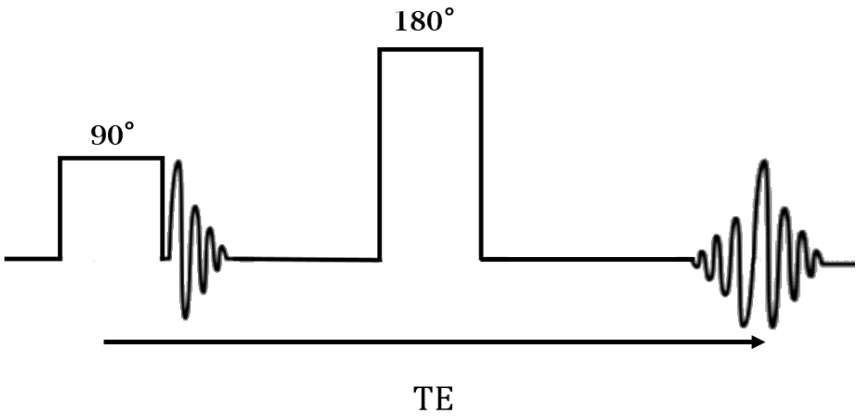

Figure 2.3. A spin echo sequence showing the initial 90° pulse followed by the 180° pulse. TE = echo time.

In a GE sequence immediately after the RF excitation pulse, a magnetic field gradient is applied, which causes the spins to rapidly dephase. Subsequently, a second gradient with opposite polarity is applied, which reverses the dephasing caused by the first applied gradient, causing the spins to come back into phase and create a gradient echo. Importantly, only the additional dephasing which was caused by the first magnetic field gradient is rephased. This can be seen in Figure 2.4.

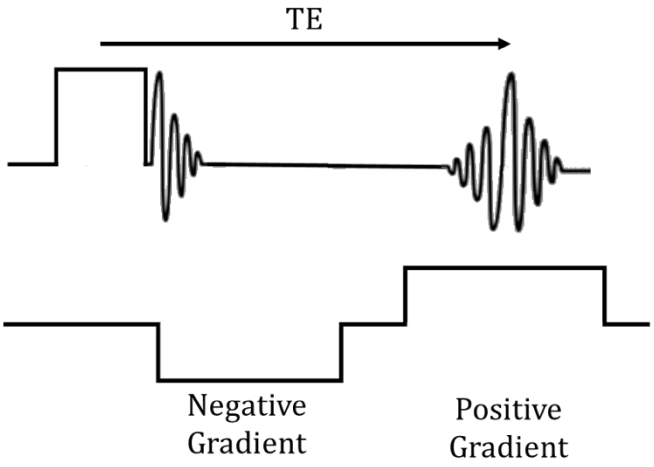

Figure 2.4. A gradient echo sequence showing the initial RF excitation pulse followed by the negative and positive magnetic field gradients. TE = echo time.

2.3: Relaxation

After the excitation pulses are turned off the spins gradually return to equilibrium. This process is known as relaxation and has two key features; the dephasing of the spins in the transverse plane and realignment of the net magnetization along the z-axis. The dephasing of the spins occurs not only due to inhomogeneity of the main magnetic field, but also because of the interactions between the spins. This interaction is known as spin-spin relaxation. This can be characterized by the spin-spin relaxation time, also known as the transverse relaxation time or T_2. The T_2 of a tissue is given by the time for the transverse magnetization to decease to 37% of its initial value.

Dephasing in the transverse plane does not result in a net loss of energy. However, as protons interact with the lattice (the surrounding tissue) energy is lost, and this is known as spin-lattice relaxation. It is characterised by the spin-lattice relaxation time, also known as the longitudinal relaxation time or T_1. The T_1 of a tissue is given by the time for the longitudinal magnetisation to recover to 63% of its equilibrium value. T_1 depends on field strength and will generally increase as field strength increases, while T_2 will decrease with higher field strength, however the T_2 effects are much smaller. T_2 and T_1 are independent processes, and the dephasing of T_2 is a quicker process than T_1 relaxation.

2.4: Contrast Agents

Contrast agents can be used to modify the T_1 and T_2 times of tissues, and thereby enhancing contrast in the resulting MR images. Gadolinium is a paramagnetic contrast agent and therefore has unpaired electrons, causing it to become temporarily magnetised when placed in a magnetic field. A gadolinium contrast agent can be injected into the bloodstream to highlight pathology and tissue status. In the healthy brain it cannot cross the blood brain barrier (BBB) and thus it is useful for highlighting lesions where the BBB is compromised. In the body it is distributed to the extracellular fluid spaces and is eventually excreted via the kidneys, and so those with kidney failure should not be administered gadolinium. The strong paramagnetic effects of gadolinium cause a decrease in local T_1 relaxation times, this results in increased signal intensity on T_1-weighted images.

2.5: Spatial Encoding

When someone is placed in an MR scanner, all the protons within their body are experiencing approximately the same magnetic field and are thus excited by the same frequency RF pulse, applied at the Larmor frequency. The 'raw' signal received is not dependent on location and therefore the signal needs to be localised. This localisation is done by applying magnetic field gradients, which can be applied in 3 orthogonal directions (along the x-axis, y-axis and z-axis).

Gradients are additional magnetic fields with linear variations which modify the strength of the main external magnetic field, B_0. During the application of these gradients, the protons are exposed to different magnetic field strengths, meaning they have different precessional frequencies (based on Equation 2.1). The slice selecting gradient (G_{ss}) (for example applied in the z-direction) and RF pulse are applied simultaneously, and the application of the slice-select gradient ensures that only spins within a selected region are precessing at frequencies which are on-resonance with the applied RF pulse. The slice thickness is determined by the bandwidth of the RF pulse, and the gradient strength.

To locate the point from which the signal originates within the selected slice, two further gradients are applied in the remaining two orthogonal directions; the frequency encoding gradient and phase encoding gradient.

The frequency encoding gradient (G_{FE}) is applied after the G_{ss} in another orthogonal direction (for example the y-axis). This causes a small, spatially-dependent

variation in the precessional frequencies along this axis, and as such the location along this axis is encoded by frequency.

The phase encoding gradient (G_{PE}) is applied along the remaining axis, in this case, the x-axis. The G_{PE} is turned on only for a short time after the RF pulse, and during this time the protons will precess with varying frequencies along the x-axis. This gradient is then turned off, and all the protons return to their original precessional frequency. However, the protons will now be out of phase with each other, and this phase change can be detected and encoded along this axis.

By combining the G_{FE} and G_{PE} following the G_{SS}, the relationship between the MR signal and the spatial location of the spins contributing to this signal can be determined within the selected slice.

2.6: Combining Gradients and Pulse Sequence Diagrams

As explained previously (Section 2.5), the G_{ss} selects the appropriate slice to be imaged, and through the subsequent addition of the G_{PE} and G_{FE} the signal can be localised. To obtain an image in MRI, the process of exciting protons, and applying RF pulses and gradients is repeated many times. This a controlled process, with specific timings, and can be represented in a pulse sequence diagram. Two key timings are often referred to, these are the TE (echo time) and TR (repetition time). TE refers to the time between the application of the RF pulse and the peak of the induced echo. TR refers to the time between the application of an RF excitation pulse and the next RF pulse. An example of a gradient echo pulse sequence diagram is shown in Figure 2.5.

Alterations to TE and TR can affect the contrast of resulting MR images. For example, a short TR and TE will produce a T_1-weighted image where CSF (cerebral spinal fluid) appears dark, however a long TR and TE will produce a T_2-weighted image where CSF appears bright.

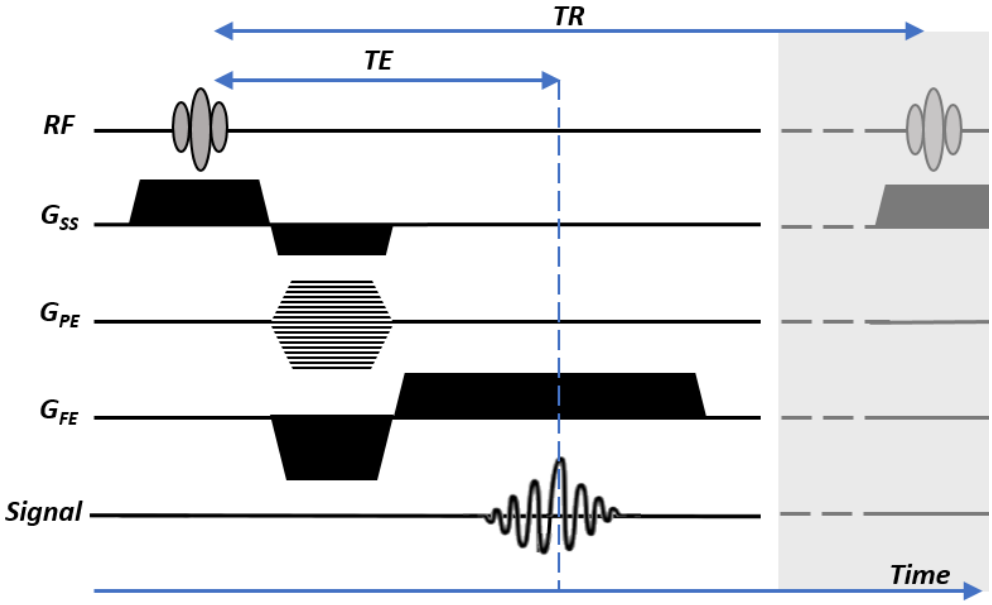

Figure 2.5. A gradient echo sequence showing the RF pulse, G_{SS} (slice selecting gradient), G_{PE} (phase encoding gradient), G_{FE} (frequency encoding gradient), with time running horizontally. The diagram also shows the TE (echo time) and TR (repetition time).

Summary of Fundamentals of MRI

This chapter provided a brief overview of the fundamental processes involved in MRI which are required to obtain an image. There are numerous sequences in MRI which can provide different contrasts and information regarding the tissue under investigation. The following chapter (Chapter 3) explores one of these techniques.

References: Chapter 2

1. Dale BM, Brown MA, Semelka RC. MRI Basic Principles and Application. 5th ed. John Wiley & Sons; 2015.
2. McRobbie DW, Moore EA, Graves MJ, Prince MR. MRI from Picture to Proton. 2nd ed. Cambridge: Cambridge University Press; 2006.
3. Plewes DB, Kucharczyk W. Physics of MRI: A primer. *J Magn Reson Imaging* **35**, 1038–54 (2012).
4. Pooley RA. Fundamentals of MR Imaging. *RadioGraphics* **25,** 1087 – 1099 (2005).
5. Schild HH. MRI Made easy 2nd ed. Berlin: Schering; 1997.
6. Naaby O. (2019). Classical and Quantum Explanations of Magnetic Resoannce. Viewed 29 March 2019, http://www.drcmr.dk/MR.
7. Stejskal EO, Tanner JE. Spin diffusion measurements: spin echoes in the presence of time-dependent field gradient. *J Chem Phys* **42,** 288-292 (1965).

Chapter 3: Diffusion Weighted Imaging and Models of Diffusion

This chapter focuses on explaining the underlying processes of diffusion weighted imaging (DWI). It also explores four models used to describe the diffusion signal; mono-exponential, IVIM (intravoxel incoherent motion), stretched exponential and kurtosis. The chapter provides an overview of each model and gives examples of the clinical utility of the models in the body. Specific details about the use of diffusion imaging in Wilms' tumour are provided in Chapter 4.

3.1.1: Background of Diffusion Weighted Imaging and the Apparent Diffusion Coefficient

Diffusion weighted imaging (DWI) is a semi-quantitative MRI technique, based on the random Brownian motion of water molecules[1]. The displacement of water molecules over a given time (τ) can be described though Einstein's equation (Equation 3.1):

$$\langle r^2 \rangle = 6D\tau \quad [3.1]$$

$\langle r^2 \rangle$ = mean square displacement

D = diffusion coefficient

When there is no internal structure to inhibit the motion of these molecules it is known as free diffusion. However, within the body there are different structures (such as cell membranes) which can hinder the motion of water molecules. For this reason, DWI can be greatly informative about tissue structure and cellular environments within the body.

A common method for achieving a diffusion weighted image is to use a pulsed gradient spin echo (PGSE) sequence[2]. A diagram for this can be seen in Figure 3.1 PGSE consists of a 90° RF pulse followed by a diffusion gradient applied along a given axis, which dephases the spins. Following this a 180° RF pulse is applied and then another diffusion gradient. The gradient moments of both the diffusion gradients are equal in magnitude.

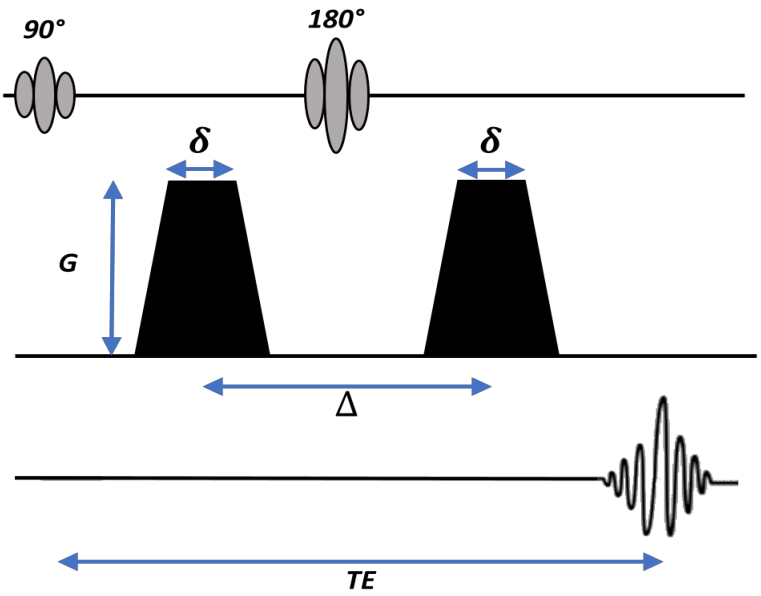

Figure 3.1. A pulsed gradient spin echo sequence showing the two RF pulses (90° and 180°) and the two diffusion gradients with magnitude G, spacing Δ and duration δ. TE = echo time.

The strength of the gradient (G), the duration (δ) and time between the gradients (Δ) control the diffusion weighting, otherwise known as the b value, which is measured in units of s/mm^2.

In a PGSE sequence b can be calculated by:

$$b = \gamma^2 G^2 \delta^2 \left(\Delta - \frac{\delta}{3}\right) \quad [3.2]$$

The parameter $\left(\Delta - \frac{\delta}{3}\right)$ is also known as the diffusion time (τ), mentioned above.

The degree of signal attenuation is proportional to the amount of diffusion in the direction of the applied diffusion gradients. The application of the first gradient will cause the spins to become out of phase and the second gradient will rephase only the static spins. For example, if there is no movement between these two gradients; if diffusion is restricted, then the spins will be become back in phase producing a high signal. Therefore, on a diffusion weighted image in areas where tissue has a higher cellularity, and diffusion is more restricted, the signal will appear bright. However, if the spins have moved between the gradients the second gradient will not cause the spins to become back in

phase and thus these areas will have a reduced signal intensity on a diffusion weighted image. The strength of the signal can be given by:

$$S(b) = S_0 e^{-bD} \quad [3.3]$$

Where S(b) is the signal intensity as a function of b value (b). S_0 is the signal intensity obtained without diffusion weighting (b=0). The parameter *D* is the diffusion coefficient, however, as this is influenced by a number of factors, such as the presence of cell membranes, we refer to this as the ADC; the Apparent Diffusion Coefficient. ADC maps can be calculated by acquiring data using two or more b values. Tissue with high ADC values will be represented as bright on an ADC map.

3.1.2: Clinical Utility of ADC in the Kidneys

ADC has been used in a variety of organs and pathologies; however, the following section focuses on research applications in the kidneys.

Previous work has highlighted that ADC in normal renal parenchyma is significantly higher than ADC in a large range of solid renal tumours[3]. This is most likely because a highly dense cellular environment, such as a malignant solid tumour, would have reduced diffusion and thus a lower ADC. Moreover, it has also been shown that benign lesions had significantly greater ADC values compared to malignant lesions[4]. This retrospective study also showed that benign cysts had significantly higher ADC values than renal cystic cancers, a finding which has also previously been reported[5]. The lower ADC values in malignant cystic tumours may be due to the increased blood and protein products in complex cysts and also larger molecules and cellular debris which could hinder diffusion[4]. The authors also suggested that ADC may be useful in identifying cysts and malignancies when gadolinium is not an option, due to patients' impaired renal function[4]. Furthermore, in a prospective study it was shown that those with renal failure had significantly lower ADC values than healthy volunteers[6]. The implication was that water was restricted in the extravascular extracellular spaces resulting in the reduction of ADC. Thus, ADC has the potential to provide clinically useful information regarding renal function.

A retrospective study of chronic kidney disease also highlighted the efficacy of ADC[7]. It was shown that median ADC was significantly greater in healthy tissue than those with stage 3, 4, or 5 chronic kidney disease. Furthermore, ADC could significantly differentiate between controls, stage 1, and stage 2 groups compared with stage 3, 4 and 5 groups with reasonably high sensitivity and specificity (75% and 69%, respectively). The authors did note that determining an ADC cut-off value to separate these stages is challenging, as ADC could be affected by b value ranges and field strength, thus a much larger study with a range of patients and protocols would be needed to provide a cut-off which could be applied clinically.

Other research has highlighted that ADC could be correlated with renal function. For example, it was shown that ADC was positively correlated with residual renal function in continuous ambulatory peritoneal dialysis patients[8]. Furthermore, it has also been shown that ADC correlated with estimated glomerular filtration rates (eGFR) in patients with diabetic nephropathy and in those with chronic kidney disease without diabetes[9]. Biopsies from the chronic kidney disease group also showed that ADC correlated with renal pathology. There is therefore potential for ADC to assess renal function as many renal processes are associated with diffusion of water, such as tubular reabsorption, secretion, and glomerular filtration[7].

Overall it can be seen that ADC has great potential to be used clinically to assess normal renal function and a range of renal impairments.

3.2.1: The IVIM Model (Intravoxel Incoherent Motion)

IVIM is a bi-exponential model of diffusion introduced by LeBihan (1988)[10]. The mono-exponential model which provides a measure of apparent diffusion of water molecules was thought to be too simplistic, thus IVIM was proposed. IVIM aims to describe the relationship of increasing b values and signal attenuation by not only describing the movement of the water within the extra-vascular / extra-cellular space, but also within the micro capillary network. IVIM produces the fitted parameters; *D, D** and *f* and is given by Equation 3.4[10]:

$$S(b) = S_0[(1-f)e^{(-b.D)} + fe^{[-b.(D+D^*)]}] \quad [3.4]$$

S(b) = signal at a given b value

S_0 = signal with no diffusion weighting

D = 'Slow' Diffusion Coefficient

*D** = 'Fast' Diffusion Coefficient

f = Perfusion Fraction

The S_0 parameter represent the signal from b = 0 and the remaining parameters represent different contributions from the diffusion signal and micro capillary perfusion[11]. *D* represents the molecular diffusion of water through static tissue and can represent hindered diffusion due to cellular structures, which is free from the influence of fast flowing water in the capillary network. *D* is therefore often referred to as 'slow' diffusion. *D** can be thought of as the 'fast' diffusion coefficient, which is due to the motion of water molecules in the blood as it flows in the randomly orientated capillary network. As water is flowing in many directions across this network, *D** is sometimes referred to as 'pseudo-diffusion', as the combination of randomly orientated flow in this network can appear like diffusion. At higher b values the contribution of 'fast' diffusion diminishes, however at lower b values (0 – 200 s/mm^2) the influence can be significant. The *f* parameter is also known as the 'perfusion fraction' which represents the intra-voxel volume fraction associated with the fast diffusion component and can act as a surrogate biomarker for the local microvascular volume fraction.

An example of an IVIM fit to raw DWI data can be seen in Figure 3.2; where a bi-exponential fit provides a good description of the raw DWI data. It can be seen in Figure 3.2 that there is initially a fast decay in signal, followed by more gradual signal attenuation (b values > 200 s/mm^2).

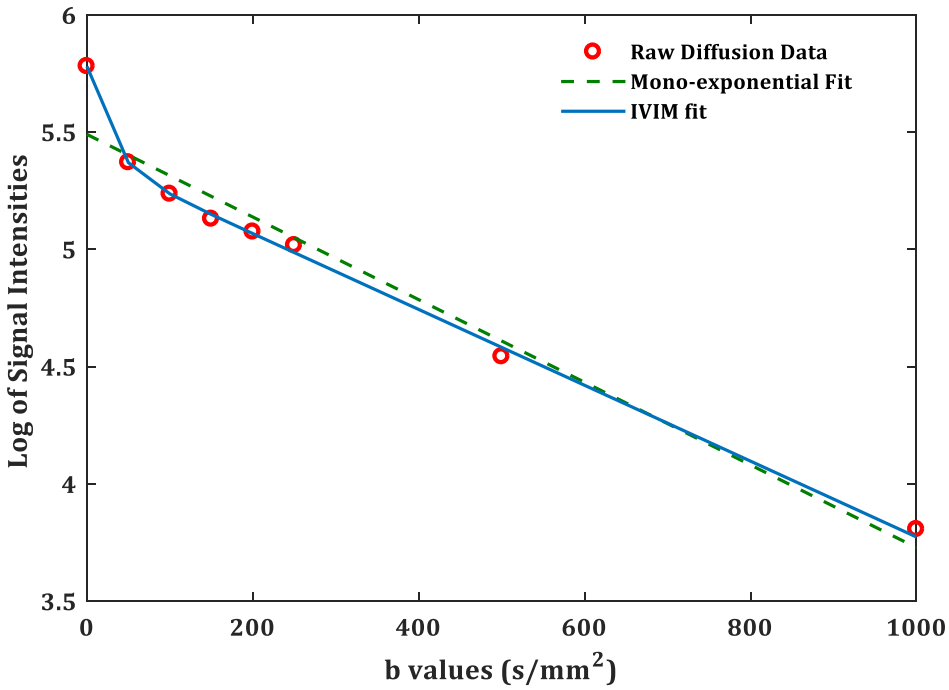

Figure 3.2. An example of the IVIM effect within a healthy adult kidney. The plot shows the natural logarithm of the signal intensities against the b values from normal kidney tissue (red circles). The IVIM fit (blue solid line) is shown against the raw data, along with a mono-exponential fit (dashed green line).
Data is based on the kidney of a healthy adult male (age 28.26 years) from a 3T scanner, with b values = 0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2

It has been noted that while *D** may represent the fast diffusion component and be sensitive to micro-capillary perfusion, it may also be influenced by other factors such as tubular flow or glandular secretion[11]. It was therefore suggested that it may be difficult to isolate perfusion effects and caution should be taken when interpreting this parameter[11].

The choice of b values is important in the IVIM model; crucially, low b values (< 200 s/mm^2) are needed to capture the perfusion effects. However, there is no universally agreed consensus on the optimal range of b values to use[11]. Using many b values over a large range may produce high quality data, however due to the increase in scan time this is not always clinically feasible.

3.2.2: Fitting Methods for the IVIM model

There are various methods for performing an IVIM fit, for example parameters can be fit in a piece-wise manner or fit simultaneously[12]. There are different forms of piece-wise fitting; for example, fitted values could be constrained by using a mono-exponential

fit at high b values (for example: ≥ 200 s/mm^2) to determine *D*, and then subsequently fitting for *D** and *f*. This is to ensure that *D* is not influenced by *D** and *f* and is possible because the influence of *D** is negligible at these higher b values. Another version of piece-wise fitting builds on the previously mentioned model by determining *D* at high b values and then calculating *f* by extrapolating a linear fit of the natural logarithm of signal intensity back to b=0 and taking the intercept. With *D* and *f* now fixed, *D** can then be fit to the raw data. Alternatively, an un-constrained approach could be used, in which all parameters are fit simultaneously.

Studies have compared these fitting methods and have shown that the most constrained method (constraining *D* and *f* prior to fitting for *D**) outperform the simultaneous[12,13] and partially constrained methods[12]; providing more reproducible and accurate (based on simulated data) diffusion parameters. However, the simultaneous and less constrained methods have been shown to provide better fits to the raw data (according to the Akaike Information Criterion [AIC]) compared to the more constrained method[14]. Although, it should be noted that the more constrained method did provide more robust and reproducible values of *D** and *f* compared to the other two methods[14]. Thus, it can be challenging to determine which is the most appropriate method, with unconstrained fitting methods providing the best fit to the raw data, but at the expense of poorer reproducibility in the values of the fitted parameters.

Other fitting methods have also been investigated; such as Bayesian fitting methods, which have also been compared to the previously mentioned most constrained technique[15]. It was shown that all methods were highly similar in terms of reproducibility and accuracy. Thus, the authors suggested that the constrained method should be favoured as it is less numerically complex and computationally demanding.

3.2.3: Clinical Utility of the IVIM Model in the Body

The IVIM model has been applied in a wide range of tissues across the body; for example, in the prostate[16,17], pancreas[18], and liver[19,20]. In a study of healthy volunteers and patients with prostate cancer it was shown that *D* and *D** values were significantly lower in cancerous tissue compared to healthy tissue[16]. Additionally, this was also seen in *f*, but only in the peripheral zone of the prostate. This is in line with previous work, where it was shown that ADC of cancerous tissue was lower than that of benign prostate tissue[21,22]. This is likely to be due to the increased cellularity in cancerous tissue, leading to a decrease in apparent diffusion. Therefore, it is unsurprising these effects are also present when using

the IVIM model. However, the IVIM model did provide a superior fit to the raw data, compared to a mono-exponential model, when compared using AIC[16]. This has also been shown in prostate cancer[17].

Research into the pancreas has highlighted the utility of the *f* parameter[18]. When those with pancreatic carcinoma were compared to the healthy volunteers it was shown that *f* was superior to ADC in terms of visual depiction of pancreatic carcinoma. Furthermore, it was also shown that *f* had a significantly larger AUC (area under the curve) compared to *D* and *D** ($p < 0.05$) when distinguishing between cancerous and healthy tissue according on ROC curves (receiver operator curves). Although it should be noted that this study used many b values, with several smaller b values (5 b values < 100 s/mm^2), and employed breath hold techniques to minimise registration errors. Thus, other studies may struggle to replicate these results.

In the liver it has been shown that *D** was able to highlight decreased perfusion in liver cirrhosis compared to healthy liver tissue[19]. Furthermore, both *D* and *D** showed significantly lower values in malignant hepatic lesions compared to benign[20].

As previously mentioned the fast diffusion coefficient, *D**, may also be related to laminar flow, and thus it is very difficult to isolate the microcapillary perfusion effects, and differentiate these from 'through-flowing' water in larger vessels[11]. Due to this, it was previously suggested that IVIM may not be able to measure vascular perfusion in the kidneys[23]. It was argued that high values of fast-diffusion in the kidneys are likely to be more related to renal tubular flow as opposed to vascular perfusion. It was claimed that even if vascular perfusion was contributing to the signal it would not be possible to isolate this effect using the IVIM model. However, while the exact physiological basis of the fitted parameters remains unclear, it does not prevent this model from being clinically useful in renal studies.

The IVIM model showed promise in detecting early changes in diabetic kidneys; with *f* being significantly higher in diabetic kidneys compared to controls and *D* being significantly lower. *D**, however, did not show significant differences[24]. Furthermore, IVIM parameters have been shown to be useful in highlighting kidney function in those with chronic kidney disease, with both *D* and *f* being related to estimated glomerular filtration rate (eGFR), again *D** was not shown to be a useful parameter[25]. Previous work has also shown that *D* was significantly related to eGFR, and additionally this work also showed that *D*, *D** and *f* were significantly lower in those with severe renal injury compared to those without[26]. However, other work has opposed this, showing that none of the IVIM

parameters were related to eGFR[27]. Furthermore, although the IVIM parameters were related to renal injury, the model did not add significantly to the assessment of renal dysfunction when compared to ADC[26].

It can be seen that it is possible to use the IVIM model in the kidneys, and that there is potential for this model to be clinically useful. However, clearly more work is needed to understand the benefits of this model and whether all the IVIM parameters are useful in renal investigations.

3.3.1: The Stretched Exponential Model

The stretched exponential model of diffusion was introduced to describe the heterogeneity in diffusion within a single voxel[28]. Initially proposed to describe the heterogeneity of diffusion measurements in the cerebral cortex[28]; the stretched exponential model provides a measure of the deviation from a mono-exponential signal decay. The model includes the fitted parameters *DDC* and α, and the parameter S_0 is the signal at b=0 and is described by Equation 3.5[28].

$$S(b) = S_0 e^{(-(b.DDC)^{\alpha})} \quad [3.5]$$

S(b) = signal at a given b value

S_0 = signal with no diffusion weighting

DDC = the Distributed Diffusion Coefficient

α =the stretching parameter describing the deviation from a mono-exponential decay

The stretching parameter (α) ranges between 0 and 1; the nearer to 1 the closer the signal decay is to a mono-exponential function. A lower value of α indicates greater inhomogeneity, suggested to be due to multiple discrete proton pools within the same voxel[28].

If b**DDC* > 1 then signal attenuation occurs with a decrease in α as a function of b. Conversely, when b**DDC* < 1, a decrease in α produces a faster decay in signal[28]. An example of this change in signal attenuation is given in Figure 3.3. At b**DDC* = 1 the lines cross, this can represent the separation between 'high' and 'low' ranges of decay rates[28]. At higher values of b**DDC* the signal attenuation is more reliant on the slow decaying terms, as the faster decay terms have already disappeared.

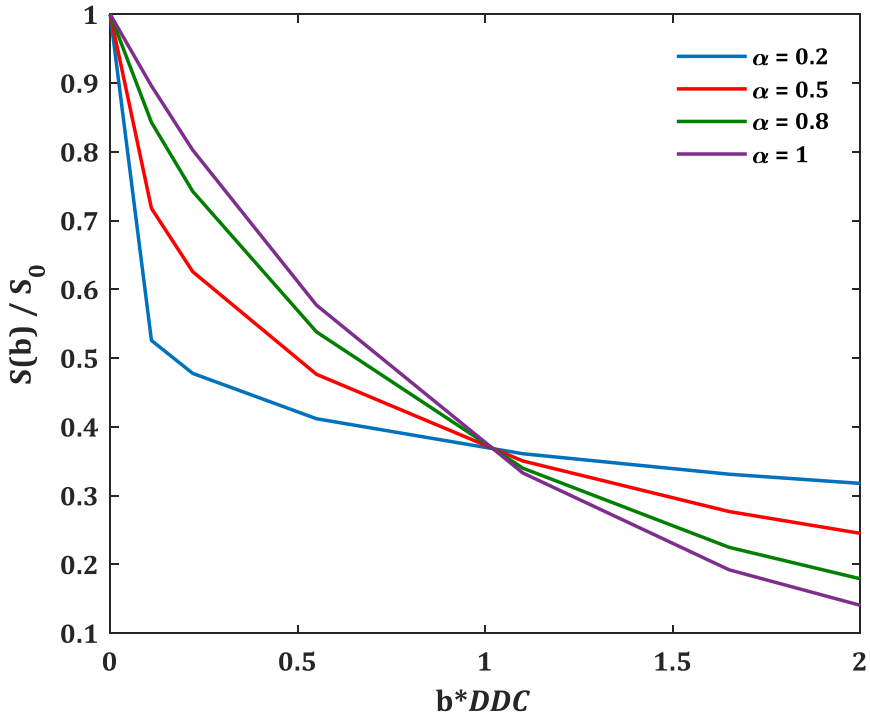

Figure 3.3. The differences in signal decay given by the stretched exponential function according to varying values of α. Signal is displayed as $S(b)/S_0$ where S(b) is the signal at a given b value and S_0 is the signal with no diffusion weighting. This signal is plotted against b**DDC* where b is the b value and *DDC* is the distributed diffusion coefficient. Data is based on *DDC* = 2.2 *10^{-3} mm/s^2, and b values = 0 50 100 250 500 750 1,000 s/mm^2.

The 'stretching' aspect of the model is observed at higher b values where lower values of α prevent the model from decaying too rapidly, however maximum b value selection will depend on the structure being imaged. Studies of the kidney[29] and bladder[30] have used a maximum b value of 1,700 s/mm^2, while in the liver[31] a value of 1,000 s/mm^2 was deemed appropriate. Furthermore, in the prostate it has been suggested that b values up to 2,000 s/mm^2 provide greater diagnostic performance in *DDC*, compared to a maximum b value of 1,000 s/mm^2 [32]. This is also in agreement with previous work in prostate cancer[33]. Additionally, it was suggested that b values > 2,000 s/mm^2 do not improve the diagnostic ability of this parameter in the prostate[32]. However, it should be noted that a maximum b value for stretched exponential modelling has not been conclusively determined in the prostate[32], or in other body organs.

3.3.2: Clinical Utility of the Stretched Exponential Model in the Body

As the stretched exponential model was originally proposed for use in the cerebral cortex research has mostly focused on its utility in the brain. For example, in gliomas the α parameter managed to significantly separate tumour from normal tissue[34] additionally α provided greater diagnostic power in differentiating higher and lower grade gliomas compared to ADC[35].

However, there is also some literature on the clinical utility of this model in the body. Here, the stretched exponential model has demonstrated that both *DDC* and α were significantly lower in prostate cancer when compared to normal prostate tissue[36]. This study also compared the stretched exponential model to a mono-exponential model. There was a strong relationship between ADC and *DDC*, particularly when measured in urine, where values were near equal. This relationship can be explained as urine is similar to pure water, thus would not deviate from a mono-exponential pattern, and therefore α would be very close to 1. Within areas of normal prostate tissue there was also high similarity between ADC and *DDC*, this was explained due to the many fluid filled structures and large areas of free water, thus also giving α values close to 1. Within the cancer tissue there was greater discrepancy between ADC and *DDC*, due to the more heterogeneous nature of the tissue, thus giving a lower value of α. The authors concluded that the stretched exponential model, and the inclusion of the extra parameter α, may add further diagnostic information when distinguishing between benign and malignant prostate tissue.

The utility of the stretched exponential model in body imaging has been further examined by comparing it to a mono-exponential and a bi-exponential model in the kidneys[29]. This study sought to determine which parameters from a variety of models were best suited for distinguishing between minimal fat angiomyolipoma and renal cell carcinoma. Compared to the other models, *DDC* was the weakest parameter, failing to reach significance to discriminate between the two tissue types. However, α reached significance (p <.001), and produced far higher levels of sensitivity (90.5%) and specificity (84.2%) compared to ADC, and D^* and f from the IVIM model.

Research in both brain and body imaging has suggested that α may be an important diagnostic parameter[29,34–36], however *DDC* may not provide as pertinent information given that it has been shown to be very similar to ADC[36] and also failed to reach significance in the previously mentioned study[29].

3.4.1: The Kurtosis Model

In the case of completely free diffusion, a mono-exponential decay of signal with increasing b value would be expected[37]. However, in reality water molecules do not always obey free diffusion due to restrictions of diffusion caused by cellular structures, therefore the kurtosis model is used to describe this deviation from a Gaussian distribution[38]. The model aims to describe the non-Gaussian behaviour of diffusion which may reflect tissue heterogeneity, and provide important information about tissue microstructure[39]. The kurtosis model produces the fitted parameters D_k and K, and the S_0 parameter is the signal at b=0, the model is described by Equation 3.6[38].

$$S(b) = S_0 e^{-b.D_k + b^2 {D_k}^2 K/6} \quad [3.6]$$

S(b) = signal at a given b value

S_0 = signal with no diffusion weighting

D_k = the diffusion coefficient corrected for the non-Gaussian displacement of water molecules

K = the Kurtosis

From a mathematical perspective, the kurtosis value describes the peaks of a curve compared to a Gaussian bell curve[40]. A kurtosis (K) of zero indicates a perfectly Gaussian distribution and a distribution where $K > 0$ would have higher peaks and sharper tails, this can be seen in Figure 3.4. $K < 0$ is possible from a mathematical perspective, however in the context of diffusion it is not possible; diffusion would have to be completely restricted which is not the case in tissue.

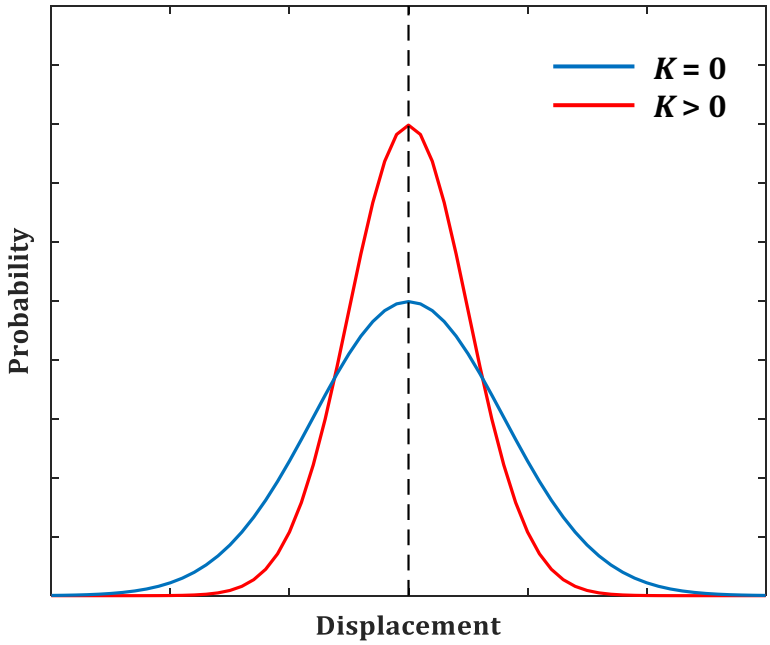

Figure 3.4. An example of a Gaussian distribution: $K = 0$ (blue line), and a non-Gaussian distribution, given by $K > 0$ (red line).

In the case of diffusion imaging, a higher value of K would suggest a more complex tissue environment[40]. Additionally, it has been suggested that K may reflect the interaction of water molecules with intracellular compounds and cell memebranes[38,41], however the exact underlying biophysical nature of K is not well defined.

The selection of b values used for the kurtosis model is very important. In a mono-exponential fit it may be reasonable to use b values up to 1,000 s/mm^2 in body imaging, however the signal deviates from a mono-exponential decay at higher b values (b > 1,500 s/mm^2), demonstrating the non-Gaussian nature of water diffusion[38]. It is thus important to use higher b values to capture the kurtosis effects. The kurtosis model becomes more sensitive to heterogeneous structures and shorter molecular distances at b values > 1,500 s/mm^2 as the signal deviates away from a mono-exponential decay[40]. Ultra-high b values (> 3,000 s/mm^2) should also be avoided when using a kurtosis fitting; rather than maintaining a decay in signal, this will cause an increase which is not physiologically accurate[42]. When selecting the upper bound of b values, it should be selected based on the rule given in Equation 3.7[42]:

$$b \leq \frac{3}{D_k^* K} \quad [3.7]$$

D_k = the diffusion coefficient corrected for the non-Gaussian displacement of water molecules

K = the Kurtosis

b = the maximum b value

Given typical values of D_k (~1 x 10^{-3} mm/s^2) and K (~ 1) in healthy brain tissue (grey matter), it has been suggested that b should not exceed 2,000 – 3,000 s/mm^2 in the brain[42]. In body imaging, however, a lower b value may be used for the upper limit as values of D_k and K vary throughout the body. For example, a maximum b value of 800 s/mm^2 has been used in the prostate[43,44] and in whole body imaging[45]. A study into the impact of b value selection on the kurtosis model, however, did highlight that higher b values (b maximum = 2,300 s/mm^2) provided the best distinction between healthy and cancerous tissue in the prostate[46].

Within the kidneys, kurtosis modelling was achieved as part of two feasibility studies with maximum b values of 600[47] and 1,000[48] s/mm^2. The upper bound of 1,000 s/mm^2 was justified by the authors as diffusion coefficients of healthy kidney tissue are around twice as high as the brain (for both cortex and medulla)[6,49], and thus based on the relationship described above (Equation 3.7), the authors hypothesised that a maximum b value of 1,000 s/mm^2 was reasonable to use. However, it has since been suggested that the deviation from a mono-exponential decay is more apparent at higher values and testing at 600 and 1,000 s/mm^2 was not an appropriate test of feasibility[50].

3.4.2: Clinical Utility of the Kurtosis Model in the Body

Clinical applications of the kurtosis model in body imaging have been mainly demonstrated in prostate cancer[46,50–52]. In a study comparing cancerous prostate tissue and normal contralateral peripheral zone it was shown that D_k was significantly reduced in cancer tissue, and K was significantly higher[51]. This may be because diffusion would be less restricted in benign tissue, resulting in a higher D_k and that cancerous tissue may have a more complex/heterogenous tissue environment thus producing a higher value of K.

This study also used ROC analysis (Receiver Operator Characteristics) to further investigate whether kurtosis parameters could separate the two tissue types[51]. Furthermore, it also investigated whether the kurtosis model improved this separation compared to ADC. It was shown that D_k had a significantly smaller AUC (Area under Curve) compared to ADC. *K* had a larger AUC than ADC, but this difference was not significant, *K* also had the highest sensitivity compared to the other two parameters. Therefore, although D_k may not add to the separation of benign peripheral zone and prostate cancer, *K* may provide added value.

Similar results were also highlighted in another study which aimed to separate healthy peripheral zone and prostate cancer using a range of diffusion models[46]. *K* had the highest AUC compared to ADC, D_k, *D* (IVIM) *D** (IVIM), and *f* (IVIM). However, this was only true when higher b values were used (ranges from 0-2,300 s/mm^2, and 0-1,800 s/mm^2), when a maximum b value of 800 s/mm^2 was used the AUC of *K* and ADC were comparable. This again highlights the need for higher b values in the kurtosis model.

Further research into the kurtosis model and prostate cancer has again shown that *K* could significantly distinguish between benign peripheral zone and prostate cancer and also that *K* was correlated with Gleason score[52]. Additionally, *K* has also been able to differentiate between high and low grade prostate cancers with greater sensitivity than ADC and D_k[50]. However, other research has not found this to be the case[53,54]; a mono-exponential model was found to be adequate for discriminating between tissue types and tumour characterisation in prostate cancer in studies using b values of up to 800[53] and 2,000[54] s/mm^2; it was shown in both instances the kurtosis model did not add to this estimation.

The clinical utility of the kurtosis model has also been shown in body imaging outside of the prostate[55,56]. In breast cancer it was shown that *K* was significantly higher in malignant tissue compared to benign and D_k was significantly reduced in malignant tissue[55]. Furthermore, both *K* and D_k were superior in terms of sensitivity and specificity at distinguishing between these tissue types compared to ADC. Additionally, *K* and D_k were both significantly correlated with the histological tumour grade.

A study has also investigated the use of kurtosis in determining viability of hypervascular hepatocellular carcinoma[56]. It was shown that *K* was significantly higher in viable tissue compared to necrotic, potentially because higher values of K may represent more complex tissue environments[40]. It was also shown that *K* was superior in this viable/necrotic distinction in terms of sensitivity, specificity, and AUC compared to ADC.

The kurtosis model has clearly been shown to be clinically useful in body imaging, and while it's use in kidney research is sparse it has been suggested that that this model may be useful in kidney fibrosis[57]. However, further research is needed to determine its clinical utility in this organ.

Summary of Diffusion Weighted Imaging and Models of Diffusion

Chapter 3 detailed Gaussian and non-Gaussian models of diffusion and reviewed research into their clinical utility in the body. The mono-exponential model, which provides ADC, has been widely applied to many organs. It does not need a large range of b values and thus is relatively easy to acquire in clinical environments.

There has been less research into non-Gaussian models of diffusion in the body. However, from reviewing the literature a few key points can be identified. The IVIM model can provide additional information regarding microcapillary perfusion as well as diffusion information and can be clinically useful in describing renal function. It also seems that constrained step-wise fitting methods of IVIM are preferable to simultaneous fittings.

The stretched exponential model describes the deviation from a mono-exponential decay, with increased deviation representing increased tissue inhomogeneity. Based on the described research *DDC* does not seem to be as clinically useful as α.

The kurtosis model describes the non-Gaussian nature of diffusion and provides two fitted parameters: D_k and K. Both D_k and K have been shown to be clinically useful, particularly in the prostate, and it has been demonstrated that this model may be superior to mono-exponential fitting when discriminating between benign and malignant tissue.

References: Chapter 3

1. Koh, D.-M. & Collins, D. J. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *Am. J. Roentgenol.* **188**, 1622–1635 (2007).
2. Stejskal, E. O. & Tanner, J. E. Spin Diffusion Measurements: Spin Echoes in the Presence of a Time-Dependent Field Gradient. *J. Chem. Phys.* **42**, 288–292 (1965).
3. Squillaci, E. *et al.* Diffusion-weighted MR imaging in the evaluation of renal tumours. *J. Exp. Clin. Cancer Res. CR* **23**, 39–45 (2004).
4. Sandrasegaran, K. *et al.* Usefulness of Diffusion-Weighted Imaging in the Evaluation of Renal Masses. *Am. J. Roentgenol.* **194**, 438–445 (2010).
5. Zhang, J. *et al.* Renal masses: characterization with diffusion-weighted MR imaging--a preliminary experience. *Radiology* **247**, 458–464 (2008).
6. Thoeny, H. C., De Keyzer, F., Oyen, R. H. & Peeters, R. R. Diffusion-weighted MR imaging of kidneys in healthy volunteers and patients with parenchymal diseases: initial experience. *Radiology* **235**, 911–917 (2005).
7. Yalçin-Şafak, K. *et al.* The relationship of ADC values of renal parenchyma with CKD stage and serum creatinine levels. *Eur. J. Radiol. Open* **3**, 8–11 (2015).
8. Yang, L., Li, X.-M., Zhao, S., Hu, Y.-J. & Liu, R.-B. Diffusion-Weighted Imaging of the Kidneys and Its Relationship With Residual Renal Function in Continuous Ambulatory Peritoneal Dialysis Patients. *Am. J. Roentgenol.* **204**, 1008–1012 (2015).
9. Inoue, T. *et al.* Noninvasive evaluation of kidney hypoxia and fibrosis using magnetic resonance imaging. *J. Am. Soc. Nephrol. JASN* **22**, 1429–1434 (2011).
10. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
11. Koh, D.-M., Collins, D. J. & Orton, M. R. Intravoxel Incoherent Motion in Body Diffusion-Weighted MRI: Reality and Challenges. *Am. J. Roentgenol.* **196**, 1351–1361 (2011).
12. Meeus, E. M. *et al.* Evaluation of intravoxel incoherent motion fitting methods in low-perfused tissue. *J. Magn. Reson. Imaging* **45**, 1325–1334 (2017).
13. Cho, G. Y. *et al.* Comparison of fitting methods and b-value sampling strategies for intravoxel incoherent motion in breast cancer. *Magn. Reson. Med.* **74**, 1077–1085 (2015).
14. Merisaari, H. *et al.* Fitting methods for intravoxel incoherent motion imaging of prostate cancer on region of interest level: Repeatability and gleason score prediction. *Magn. Reson. Med.* **77**, 1249–1264 (2017).

15. Jalnefjord, O. *et al.* Comparison of methods for estimation of the intravoxel incoherent motion (IVIM) diffusion coefficient (D) and perfusion fraction (f). *Magn. Reson. Mater. Phys. Biol. Med.* (2018). doi:10.1007/s10334-018-0697-5
16. Shinmoto, H. *et al.* Biexponential apparent diffusion coefficients in prostate cancer. *Magn. Reson. Imaging* **27**, 355–359 (2009).
17. Mulkern, R. V. *et al.* Biexponential characterization of prostate tissue water diffusion decay curves over an extended b-factor range. *Magn. Reson. Imaging* **24**, 563–568 (2006).
18. Lemke, A. *et al.* Differentiation of pancreas carcinoma from healthy pancreatic tissue using multiple b-values: comparison of apparent diffusion coefficient and intravoxel incoherent motion derived parameters. *Invest. Radiol.* **44**, 769–775 (2009).
19. Luciani, A. *et al.* Liver cirrhosis: intravoxel incoherent motion MR imaging--pilot study. *Radiology* **249**, 891–899 (2008).
20. Wang, M. *et al.* Evaluation of Hepatic Tumors Using Intravoxel Incoherent Motion Diffusion-Weighted MRI. *Med. Sci. Monit. Int. Med. J. Exp. Clin. Res.* **22**, 702–709 (2016).
21. Issa, B. In vivo measurement of the apparent diffusion coefficient in normal and malignant prostatic tissues using echo-planar imaging. *J. Magn. Reson. Imaging* **16**, 196–200 (2002).
22. Sato, C. *et al.* Differentiation of noncancerous tissue and cancer lesions by apparent diffusion coefficient values in transition and peripheral zones of the prostate. *J. Magn. Reson. Imaging* **21**, 258–262 (2005).
23. Müller, M. F., Prasad, P. V. & Edelman, R. R. Can the IVIM model be used for renal perfusion imaging? *Eur. J. Radiol.* **26**, 297–303 (1998).
24. Deng, Y. *et al.* Use of intravoxel incoherent motion diffusion-weighted imaging to detect early changes in diabetic kidneys. *Abdom. Radiol.* 1–6 (2018). doi:10.1007/s00261-018-1521-4
25. Mao, W. *et al.* Chronic kidney disease: Pathological and functional evaluation with intravoxel incoherent motion diffusion-weighted imaging. *J. Magn. Reson. Imaging JMRI* **47**, 1251–1259 (2018).
26. Ding, J. *et al.* Assessment of renal dysfunction with diffusion-weighted imaging: comparing intra-voxel incoherent motion (IVIM) with a mono-exponential model. *Acta Radiol. Stockh. Swed. 1987* **57**, 507–512 (2016).

27. Bane, O. *et al.* Assessment of renal function using intravoxel incoherent motion diffusion-weighted imaging and dynamic contrast-enhanced MRI. *J. Magn. Reson. Imaging JMRI* **44**, 317–326 (2016).
28. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
29. Li, H. *et al.* Monoexponential, biexponential, and stretched exponential diffusion-weighted imaging models: Quantitative biomarkers for differentiating renal clear cell carcinoma and minimal fat angiomyolipoma. *J. Magn. Reson. Imaging JMRI* **46**, 240–247 (2017).
30. Wang, Y. *et al.* Comparison of the Diagnostic Value of Monoexponential, Biexponential, and Stretched Exponential Diffusion-weighted MRI in Differentiating Tumor Stage and Histological Grade of Bladder Cancer. *Acad. Radiol.* (2018). doi:10.1016/j.acra.2018.04.016
31. Hu, Y. *et al.* Assessment of different mathematical models for diffusion-weighted imaging as quantitative biomarkers for differentiating benign from malignant solid hepatic lesions. *Cancer Med.* (2018). doi:10.1002/cam4.1535
32. Feng, Z. *et al.* Evaluation of different mathematical models and different b-value ranges of diffusion-weighted imaging in peripheral zone prostate cancer detection using b-value up to 4500 s/mm2. *PLOS ONE* **12**, e0172127 (2017).
33. Merisaari, H. & Jambor, I. Optimization of b-value distribution for four mathematical models of prostate cancer diffusion-weighted imaging using b values up to 2000 s/mm2: Simulation and repeatability study. *Magn. Reson. Med.* **73**, 1954–1969 (2015).
34. Kwee, T. C. *et al.* Intravoxel water diffusion heterogeneity imaging of human high-grade gliomas. *NMR Biomed.* **23**, 179–187 (2010).
35. Bai, Y. *et al.* Grading of Gliomas by Using Monoexponential, Biexponential, and Stretched Exponential Diffusion-weighted MR Imaging and Diffusion Kurtosis MR Imaging. *Radiology* **278**, 496–504 (2016).
36. Liu, X., Zhou, L., Peng, W., Wang, H. & Zhang, Y. Comparison of stretched-Exponential and monoexponential model diffusion-Weighted imaging in prostate cancer and normal tissues. *J. Magn. Reson. Imaging JMRI* **42**, 1078–1085 (2015).
37. Rosenkrantz, A. B. *et al.* Body diffusion kurtosis imaging: Basic principles, applications, and considerations for clinical practice. *J. Magn. Reson. Imaging* **42**, 1190–1202 (2015).

38. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
39. Wu, E. X. & Cheung, M. M. MR diffusion kurtosis imaging for neural tissue characterization. *NMR Biomed.* **23**, 836–848 (2010).
40. Steven, A. J., Zhuo, J. & Melhem, E. R. Diffusion Kurtosis Imaging: An Emerging Technique for Evaluating the Microstructural Environment of the Brain. *Am. J. Roentgenol.* **202**, W26–W33 (2013).
41. Le Bihan, D. Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology* **268**, 318–322 (2013).
42. Jensen, J. H. & Helpern, J. A. MRI Quantification of Non-Gaussian Water Diffusion by Kurtosis Analysis. *NMR Biomed.* **23**, 698–710 (2010).
43. Quentin, M. *et al.* Comparison of different mathematical models of diffusion-weighted prostate MR imaging. *Magn. Reson. Imaging* **30**, 1468–1474 (2012).
44. Quentin, M. *et al.* Feasibility of diffusional kurtosis tensor imaging in prostate MRI for the assessment of prostate cancer: preliminary results. *Magn. Reson. Imaging* **32**, 880–885 (2014).
45. Filli, L. *et al.* Whole-body diffusion kurtosis imaging: initial experience on non-Gaussian diffusion in various organs. *Invest. Radiol.* **49**, 773–778 (2014).
46. Mazzoni, L. N. *et al.* Diffusion-weighted signal models in healthy and cancerous peripheral prostate tissues: Comparison of outcomes obtained at different b-values. *J. Magn. Reson. Imaging* **39**, 512–518 (2014).
47. Pentang, G. *et al.* Diffusion kurtosis imaging of the human kidney: a feasibility study. *Magn. Reson. Imaging* **32**, 413–420 (2014).
48. Huang, Y. *et al.* MRI quantification of non-Gaussian water diffusion in normal human kidney: a diffusional kurtosis imaging study. *NMR Biomed.* **28**, 154–161 (2015).
49. Ries, M., Jones, R. A., Basseau, F., Moonen, C. T. & Grenier, N. Diffusion tensor MRI of the human kidney. *J. Magn. Reson. Imaging JMRI* **14**, 42–49 (2001).
50. Rosenkrantz, A. B. *et al.* Assessment of hepatocellular carcinoma using apparent diffusion coefficient and diffusion kurtosis indices: preliminary experience in fresh liver explants. *Magn. Reson. Imaging* **30**, 1534–1540 (2012).
51. Tamura, C. *et al.* Diffusion kurtosis imaging study of prostate cancer: Preliminary findings. *J. Magn. Reson. Imaging* **40**, 723–729 (2014).
52. Suo, S. *et al.* Non-Gaussian water diffusion kurtosis imaging of prostate cancer. *Magn. Reson. Imaging* **32**, 421–427 (2014).

53. Roethke, M. C. *et al.* Evaluation of Diffusion Kurtosis Imaging Versus Standard Diffusion Imaging for Detection and Grading of Peripheral Zone Prostate Cancer. *Invest. Radiol.* **50**, 483–489 (2015).
54. Toivonen, J. *et al.* Mathematical models for diffusion-weighted imaging of prostate cancer using b values up to 2000 s/mm(2) : correlation with Gleason score and repeatability of region of interest analysis. *Magn. Reson. Med.* **74**, 1116–1124 (2015).
55. Sun, K. *et al.* Breast Cancer: Diffusion Kurtosis MR Imaging-Diagnostic Accuracy and Correlation with Clinical-Pathologic Factors. *Radiology* **277**, 46–55 (2015).
56. Goshima, S. *et al.* Diffusion kurtosis imaging to assess response to treatment in hypervascular hepatocellular carcinoma. *AJR Am. J. Roentgenol.* **204**, W543-549 (2015).
57. Kjølby, B. F. *et al.* Fast diffusion kurtosis imaging of fibrotic mouse kidneys. *NMR Biomed.* **29**, 1709–1719 (2016).

Chapter 4: MRI in Wilms' Tumour

This chapter explores the use of MRI specifically in Wilms' tumour. It reviews past research which has looked at differentiating between benign and malignant tumours, identifying Wilms' tumour from other abdominal tumours, changes due to chemotherapy, and subtype separation.

4.1: Background of MRI in Wilms' Tumour

MRI has been used in Wilms' tumour for many years. Some of the earliest work dates back to 1986, where the use of T_1-weighted (T_1w) and T_2-weighted (T_2w) imaging was evaluated in Wilms' tumour[1,2]. In general, Wilms' tumour appears hypointense on T_1w images, and hyperintense on T_2w images. After administration of gadolinium-based contrast agents Wilms' tumour appears heterogeneous on T_1w images, with areas of necrosis appearing darker compared to more viable tumour tissue.

Belt et al. investigated the use of MRI (0.15T) in 14 Wilms' tumour patients; 12 of whom had not undergone any treatment or surgery and 2 who were post-surgery[1]. This early work showed the potential of using MRI in Wilms' tumour. In all of the pre-surgery patients, MRI accurately identified the primary tumours and their origins. Furthermore, in 9 of the 12 pre-surgery patients, the tumour margins were also well defined. Additional research from the same time period evaluated 4 children with Wilms' tumour using a 0.3T scanner[2]. This work also showed that T_1w imaging could provide excellent depiction of the Wilms' tumour anatomy, and also identify its origins. The authors highlighted that T_1w imaging also provided the ability to clearly investigate the contralateral unaffected kidney. Both works showed the promise of MRI and suggested that in the future MRI may be able to distinguish Wilms' tumour from other solid abdominal tumours, and also may aid in staging, and reviewing treatment for these patients.

Following on from this in 1993, 12 Wilms' tumour patients were evaluated using a 1.5T scanner[3]. This work highlighted that T_2w imaging was able to distinguish between active nephrogenic rests and inactive nephrogenic rests, and also between Wilms' tumour and inactive nephrogenic rests. This was due to the fact that both active nephrogenic rests and Wilms' tumour appear hyperintense on T_2w imaging, whereas inactive nephrogenic rests appear hypointense. Furthermore, the authors found that on post-contrast

gadolinium-enhanced T_1w images Wilms' tumours were heterogeneous, whereas nephrogenic rests appeared more homogeneous.

4.2.1: Diffusion-Weighed Imaging in Wilms' Tumour: Benign and Malignant Tumour Separation

DWI has the potential to be very useful in the separation of benign and malignant tumours. For example, solid malignant tumours have higher cellularity and an increased number of cell membranes per unit volume, thus these tumours are more likely to have restricted diffusion[4] . However, benign tumours have higher water content, decreased cellularity and increased diffusion[4].

Previous research utilised DWI to determine whether ADC values could distinguish between benign and malignant solid abdominal paediatric tumours[5]. Additionally, the authors investigated whether there was a link between ADC and cellularity. Nineteen patients were included; 8 with benign tumours and 11 with malignant (3 of which were Wilms' tumour), all patients were scanned on a 1.5T scanner prior to receiving any treatment. While ADC values were generally higher in the benign patient group, the difference was not statistically significant (p = 0.069). However, a significant relationship was found between ADC and cell count; with an increase in cell count being related to a decrease in ADC (R^2 = 0.541, p <0.001). The lack of a significant difference between malignant and benign tumours may be attributed to the broad range of tumours included and the small cohort size.

More recently another study investigated whether ADC could separate malignant from benign paediatric solid abdominal tumours[6]. This study had a much larger sample size; 38 patients with 39 benign tumours, and 30 children with 34 malignant tumours (7 of which were Wilms'). This study found that there was a significant difference in ADC values between benign and malignant tumours, with benign tumours having an increased ADC. There were also good-to-excellent inter-observer correlations. Furthermore, ROC analysis showed that using a threshold of ADC: 1.29×10^{-3} mm^2/s provided high levels of sensitivity and specificity for this separation: 77% and 82% respectively. These results were based on all the tumours and when the analysis was limited to Wilms' and benign tumours, the sensitivity and specificity levels were higher. While this study indicated the benefits of ADC in identifying malignant tumours, the ADC threshold may need be taken with caution. A range of tumours with different cellular compositions from different abdominal organs were included, with each group having a small number of samples. To

ascribe a single threshold value may be premature as each of these tumour types are likely to have a range of ADC values, for example based on subtypes and presence of necrosis. Furthermore, only a few slices were sampled, not full tumour coverage, and so to give a tumour an ADC value may not be representative of the whole tumour environment, particular in the case of heterogeneous tumours such as Wilms'. With this in mind, one could conclude that ADC is useful in suggesting malignancy, but a more accurate threshold would need to be determined based on individual tumour types using full tumour coverage.

4.2.2: Diffusion Weighed Imaging in Wilms' Tumour: Separating Wilms' Tumour from Neuroblastoma

Neuroblastoma is a common paediatric tumour which can be found in the abdomen. Alongside Wilms', these two are the most common paediatric abdominal malignant tumours of adjacent organs[7]. As neuroblastoma can display radiological findings which are similar to Wilms' tumour, an investigation was carried out using ADC to distinguish between these two tumour types[8].

A retrospective study included 17 patients; 7 Wilms' tumours, and 10 neuroblastomas; this diagnosis was confirmed post surgery[8]. Each patient was scanned prior to chemotherapy on a 1.5T scanner. Two independent observers found that mean ADC values (based on a collection of ROIs spread across the tumour, avoiding cystic and necrotic regions) were significantly greater in Wilms' tumour compared to neuroblastoma. Furthermore, an ADC threshold of 0.65×10^{-3} mm^2/s was deemed to have high sensitivity (80%) and specificity (100%) for this separation. Differences were not observed between the two tumours when based on T_1w and T_2w signal intensities. Therefore, DWI shows promise as an important tool in clinical abdominal imaging. The difference in ADC values may have been due to increased levels of necrosis in Wilms' tumour and a lower cellular density compared to neuroblastomas. The inclusion of two observers and excellent agreement in ADC measurements between them (intraclass correlation coefficient: 0.955) increases the reliability of these results and thus further supporting the use of ADC in abdominal imaging.

4.2.3: Diffusion-Weighted Imaging in Wilms' Tumour: Chemotherapy Changes

DWI has also been used to investigate chemotherapy-induced changes in Wilms' tumour. While in the past a change in tumour volume post-chemotherapy may have indicated response to therapy, this is now thought to be unreliable due to the presence of necrosis; for example, a tumour may have responded to chemotherapy resulting in high levels of necrotic tissue but not have undergone vast changes in volume[9]. Furthermore, although Wilms' tumours may reduce in size, high volumes of malignant cells may still remain[10,11]. Given that a decrease in cellularity is indicative of chemotherapy response, and ADC is correlated with cellularity[5], it was hypothesised that positive shifts in ADC distributions may be reflective of chemotherapy response[9].

In a prospective study of 7 children with 9 abdominal tumours (6 of which were Wilms' tumour), ADC distributions were compared pre- and post-chemotherapy[9]. As necrotic tissue would theoretically have higher ADC values, due to the decrease in cellularity, necrotic regions were excluded from analysis. This was achieved by using T_1w imaging pre- and post-administration of gadolinium contrast. Pre-contrast enhanced T_1w images were subtracted from post-contrast T_1w images, and voxels which displayed intensities less than the mean values of the erector spinae muscles were classified as necrotic. Viable masks were then used on the ADC images to calculate median values and distributions of ADC pre- and post-chemotherapy.

The ADC distributions post-treatment deviated from a normal distribution for all the Wilms' tumour patients (p<0.001), and the median ADC values increased post-chemotherapy. There was also a wider range of ADC values following therapy compared to pre-treatment. One of the Wilms' tumours more than doubled in size after chemotherapy, but the median ADC also increased, and the distribution widened. If tumour volume was relied upon as the only measure of treatment response, this tumour would have been classed as progressive. However, given the increase in ADC, this patient did have some response to chemotherapy due to the decrease in cellularity. Two other Wilms' tumours (bilateral tumours from a single patient) showed a marked decrease in tumour volume (62% and 83%) but had only a very small increase in ADC and a high volume of viable tumour tissue remained (around 95%). This ADC pattern suggested that perhaps tumour tissue cellularity was unchanged, thus chemotherapy may not have had as much of an affect as suggested by the change in size. It should be noted that this study was conducted as a feasibility investigation and thus the sample size was very small and definitive

conclusions cannot be made, however it does highlight the potential benefits of using ADC as opposed to volume change to monitor chemotherapy response.

The precision of measuring ADC values to detect chemotherapy changes has also been investigated[12]. In a retrospective study, 13 pre-treatment and 9 post-treatment Wilms' tumours were analysed. Three readers drew ROIs and calculated the median ADC for each tumour; they were blind to each other's responses and to the histopathology reports of the patients. Additionally, one of the readers provided measurements twice, with a minimum interval of 4 weeks. Areas of necrosis were excluded using the method described previously[9]. Median ADC values had very low intra- and inter-observer variability; both had no suggestion of bias. The shift in distributions of ADC values following chemotherapy was also analysed, and again showed minimal intra- and inter-observer variability. A benefit of this study is that whole tumours were sampled, whereas often only single slice or a small collection of ROIs are analysed. As Wilms' tumours are heterogenous the single slice approach is flawed. Overall one can conclude that not only may ADC be useful in identifying chemotherapy-induced changes, but that this method is also potentially reliable.

Littooij et al[13] investigated the value of excluding necrotic regions. Median ADC values from whole tumour volumes were compared to those where necrotic tissue had been removed (using the method of gadolinium-enhanced T_1w). It was shown that in Wilms' tumours which had not undergone chemotherapy, median ADC values generally did not differ whether necrosis was included or not. However, in post-chemotherapy patients the difference was more marked. This may be because of the increased presence of necrosis within the tumour as a response to chemotherapy, resulting in a decrease in cellularity and increase in ADC in parts of the tumour[5]. Therefore, the ADC median would be affected due to the inclusion of this necrotic tissue. The method of identifying necrotic tissue is reliant on gadolinium-enhancement and thus is it difficult to conclude that truly necrotic regions were excluded due to the lack of histology. However, obtaining full coverage histology is challenging and in most clinical settings only a single section would be evaluated, therefore imaging the entire tumour and using gadolinium as a proxy for necrosis identification is a reasonable method.

4.2.4: Diffusion-Weighted Imaging in Wilms' Tumour: Chemotherapy and Subtypes

Identifying necrotic regions post chemotherapy is clearly important, and while enhancement on post-gadolinium T_1-weighted images represents a useful method of doing this, not all Wilms' tumour patients will receive gadolinium. A more recent study used the gadolinium-based method[9] to generate necrotic and viable tumour masks and then transferred these to ADC maps in 15 Wilms' tumours[14]. They then performed ROC analysis to identify an ADC threshold which best separated viable and necrotic tissue. The threshold was identified as 1.81×10^{-3} mm^2/s which gave a sensitivity of 67%, specificity of 75% and an AUC of 0.72. This threshold then allowed necrotic and viable tissue to be identified in those tumours which did not receive gadolinium. Given the potential problems with gadolinium (previously discussed in Section 1.5) this method enables a wider range of patients to be included and to undergo ADC analysis. However, it is important to keep in mind that an ADC threshold may not be adequate to separate viable and necrotic tumour tissue, as low ADC values may reflect high cellular environments but actually be necrotic due to coagulation.

After establishing the threshold, Hales et al.[14] went on to investigate chemotherapy changes with regards to Wilms' tumour subtypes. Twenty-three Wilms' tumours with ADC maps pre- and post-chemotherapy were analysed via multi-Gaussian model fitting to histograms of the ADC distribution within a given tumour. An example was shown of a blastemal Wilms' tumour (Figure 4.1.); prior to chemotherapy, the tumour had a low ADC subpopulation within the histogram, thought to represent a highly cellular environment, and thus more viable tumour tissue. There was also a small peak at a higher ADC value (above that of the previously established ADC threshold), this was thought to represent necrotic tissue. A third population was shown in between these two areas with a broad distribution, potentially representing tissue with a mixture of viable and oedematous tissue. Following chemotherapy, the high ADC peak remained in the same location with a similar peak height, the middle distribution shifted slightly to the right, possibly representing an increase in non-viable components. The lower ADC peak remained at the same value but greatly reduced in height. This suggested that necrotic / oedematous components of the tumour remained relatively unchanged by chemotherapy, whereas the volume of malignant (low-ADC) tissue was reduced.

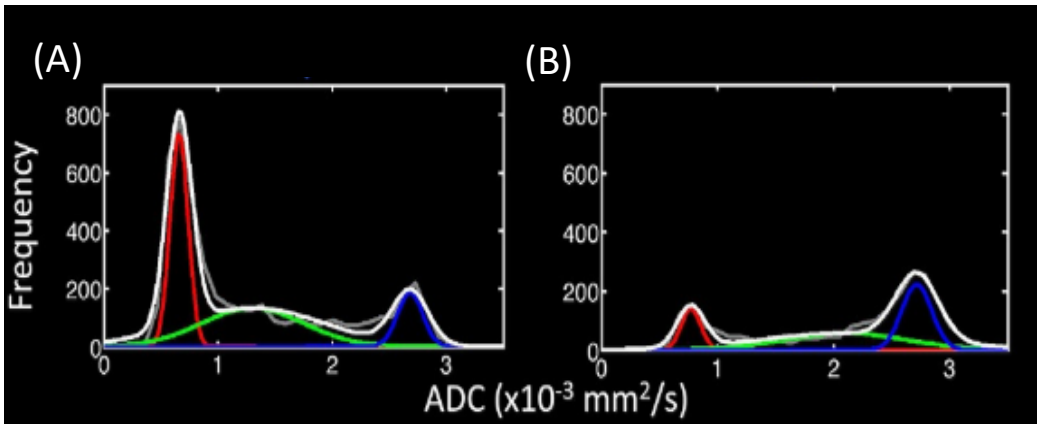

Figure 4.1. An example of ADC values in a blastemal Wilms' tumour prior to chemotherapy **(A)**, and post chemotherapy **(B)**. The different colours highlight the different underlying sub-populations in the distribution of ADC values. Taken from Hales et al *NMR Biomed*. 28, 948–957 (2015) with permission.

The study went on to explore the different effects of chemotherapy on Wilms' tumour subtypes (with the following patient numbers: stromal: 6, epithelial: 4, blastemal: 4, mixed: 7, regressive: 2)[14]. Using the mean ADC value of the highest peak in the ADC histogram (excluding values above the necrotic threshold) it was shown that only the stromal subtype significantly increased in ADC following chemotherapy.

Outside of Europe, Wilms' tumours can be treated by immediate surgery, with no prior chemotherapy. It has been suggested that chemotherapy may cause differentiation and maturation of Wilms' tumour tissue resulting in more stromal subtypes[15]. This was suggested as the incidence of stromal subtypes is very low in those who do not receive chemotherapy prior to surgery[16,17]. In the study by Hales[14] it was shown that all stromal subtypes increased in ADC following chemotherapy. This may be because prior to treatment this subtype has a high cellular density, potentially being immature blastema or stroma, and the cellular density decreases post-chemotherapy reflecting a shift to mature stromal tissue.

Blastemal subtypes have previously been noted to be the most responsive to chemotherapy[18], and if blastemal tissue remains following treatment then these patients are more likely to relapse[15]. In Hales' study[14] it was shown that the blastemal subtype decreased in volume following chemotherapy, demonstrating the responsiveness to treatment. However, the remaining chemotherapy-resistant tissue within the viable tumour retains low ADC values, indicating its high cellularity.

Hales' study also investigated whether ADC values could distinguish between the Wilms' tumour subtypes[14]. Using post-chemotherapy data from 31 Wilms' tumours (stromal: 6, epithelial: 4, blastemal: 5 mixed: 11, regressive: 5) it was shown that the previously mentioned ADC mean (highest histogram peak excluding necrosis), and the lower quartile ADC value could significantly separate blastemal from stromal, blastemal from regressive and blastemal from mixed; as the blastemal subtypes had significantly lower ADC values. When ADC of the raw data was used (not excluding necrotic tissue) these differences disappeared. This again highlights how necrosis can affect ADC values.

Littooij used histopathology to further analyse Wilms' tumour and the blastemal subtype[13]. Firstly, there was a significant inverse relationship between the amount of enhancing tissue (as assessed via T_1w gadolinium contrast) and chemotherapy–induced changes assessed by histology. The relationship between chemotherapy changes and enhancement are often assumed, and non-enhancing tissue is categorised as necrotic and removed from ADC analysis, as mentioned previously[9,12,14]. However, this study correlated the enhancement with histology, thus increasing the validity and strengthening the interpretation of this method. Furthermore, ADC was once again shown to be lower prior to chemotherapy (based on 25 Wilms' tumours with post-chemotherapy ADC maps and histology). Thus, once more highlighting the link between ADC, cellularity and chemotherapy-induced changes.

The number of each subtype was very small and so analysis of subtype separation based on diffusion measurements was not possible. Nevertheless, the proportion of blastema in 54 lesions prior to chemotherapy was weakly inversely correlated ($R = -0.248$, $p = 0.026$) with the 25th percentile of ADC (prior to chemotherapy), however the proportion of blastema in the 25 lesions post-chemotherapy did not correlate with ADC ($p = 0.072$). It is interesting that the authors attempted to measure the relationship between proportion of blastema and ADC in pre-chemotherapy tumours, as whole tumour histology is not performed until after chemotherapy. Therefore, conclusions drawn between histological subtypes and diffusion measurements prior to treatment are limited and one could argue that emphasis should be placed on the post-chemotherapy results instead.

The study also found that the stromal portion of the tumours was correlated with median ADC both pre- and post-chemotherapy[13]. Additionally, median ADC values after chemotherapy in lesions in which the viable tumour tissue was predominantly stromal were significantly higher than the median ADC values in other tumours. Using ROC analysis, an ADC value of 1.362×10^{-3} mm^2/s was able to differentiate between stromal and other subtypes with a sensitivity of 100%, specificity of 78.9%, and an AUC of 0.895.

Therefore, not only is using the viable parts of Wilms' tumour superior to using whole tumour values, but also determining stromal subtypes using ADC may aid treatment decisions. The authors suggested that the 25th percentile of ADC was the most promising metric for identifying the proportion of blastema; a finding which is similar to that of Hales[14]. However, as previously mentioned this was only for the pre-chemotherapy tumours and even this correlation was weak ($R = -0.248$, $p = 0.026$). Furthermore, the number of post-chemotherapy blastemal tumours in the cohort was very small ($n = 2$) and so despite the authors claims it cannot be said whether this marker is clinically useful. Additionally, as it has been previously shown that both blastemal and epithelial tumours have low ADC values and there is considerable overlap in ADC values between these subtypes[14], relying solely on ADC is not enough to clinically identify blastemal predominance[13].

Another study by Littooji[19] tested the feasibility of correlating DWI with post-resection histopathology, to build on their previous work and that of Hales[14]. Pathological slices were matched to T_2w or post-contrast T_1w images by a radiologist and a pathologist. Median and 25th percentile ADC values were calculated for each slice after the exclusion of necrosis. ADC values were then correlated with the percent of subtypes in the slices. Nine Wilms' tumours from 8 patients were included (mixed: 4, stromal: 2, regressive: 1, blastemal: 2). It was shown that there was a strong inverse relationship between the percent of blastema and 25th percentile ADC. Additionally, median ADC was strongly positively correlated with percent stroma. These results echo previous work indicating that blastemal tissue has lower ADC values and stromal has higher[13,14]. Furthermore, results indicated that tumour shrinkage or necrosis, in combination with low ADC values in viable tissue, was suggestive of chemotherapy-resistant blastemal tissue. This is an interesting result as remaining blastemal tissue following chemotherapy is related to increased relapse[15]. Thus, ADC values and necrotic volume change may be useful in identifying this treatment-resistant tissue. While this work suggests that there is potential to match histological slices with MRI data which may aid future work in terms of using ADC to differentiate subtypes, the lack of epithelial tissue should be noted. In this cohort none of the tumours were classified as epithelial and there was a very low proportion of epithelial tissue in each tumour. Therefore, one cannot conclude that ADC would be able to identify and separate epithelial from blastemal tissue, which is a consistent problem in subtype identification[13,14].

4.3: Diffusion Weighed Imaging in Wilms' Tumour: Non-Gaussian Models

There is a lack of research into using more advanced non-Gaussian models of diffusion in Wilms' tumour. However, one study investigated the use of IVIM in paediatric solid abdominal masses, which included 8 Wilms' tumours[20].

The retrospective study included 6 b values in their DWI sequence: 0, 50, 100, 150, 600, 1000 s/mm^2 and employed the IVIM model. The IVIM model was fit using a step-wise approach where b values above 100 s/mm^2 were used for calculation of *D*, and *f* was then calculated by extrapolating a linear fit back to b=0 and taking the intercept, following this *D** was fit. A large cohort of 42 tumours were included (10 benign, 32 malignant). The study highlighted that Wilms' tumours (n = 8) could be distinguished from neuroblastoma (n = 11) using the IVIM model.

Neuroblastomas had a higher *D** and *f* compared to Wilms' tumours. As previously suggested Wilms' tumours may have increased necrosis compared to neuroblastomas[8]. Necrotic tissue would have lower perfusion, and thus lower *f* values, which may explain the differences between these two tumours. *D** and *f* were also further investigated by conducting ROC analysis to determine whether a threshold could separate these tumour types. The mean, median, 75th and 90th percentiles of *D** all had an AUC above 0.9 for this separation. For *f*, the mean, median, 5th, 25th, 75th and 90th percentiles also all had high AUCs, with all but the 5th percentile being above 0.9 (5th percentile AUC: 0.784).

However, some results from this study were at odds with the previously mentioned work by Aslan et al.[8]; wherein ADC values in neuroblastoma were found to be significantly lower than in Wilms' tumour. The IVIM study[20] found that there were no significant differences between Wilms' tumour and neuroblastoma using ADC, or *D*. The IVIM study included the whole tumour, including areas of necrosis and cystic tissue, however Aslan[8] excluded these regions. Therefore, the lack of significant differences in the IVIM study maybe due to the inclusion of these regions which would have distorted the ADC and *D* values.

Summary of MRI in Wilms' Tumour

MRI has been used in Wilms' tumour for many years and there is currently a considerable amount of literature using DWI. DWI has been used to identify malignant from benign tumours, separate Wilms' from other abdominal tumours, and to investigate changes in chemotherapy and identify Wilms' tumour subtypes. While cohort numbers

have been relatively small, particularly when looking at subtype separation, the general trend has shown that ADC in Wilms' tumour increases following chemotherapy and excluding areas of necrosis is important to obtain accurate measurements of viable tumour. Furthermore, separating subtypes using ADC has been shown to be possible. Stromal subtypes have higher ADC values whereas blastemal have reduced ADC compared to other subtypes, apart from epithelial where there is considerable overlap in values. Additionally, matching histology to MRI slices may also be feasible. Taking into account all of the past research, diffusion imaging is a valuable tool in investigating Wilms' tumour biology and response to treatment. Furthermore, the use of non-Gaussian models may provide further information regarding the diagnosis, and assessment of treatment response of this paediatric tumour. By further investigating DWI and more advanced models, identification of tissue status, such as necrosis, and subtypes of Wilms' tumour may be possible.

References: Chapter 4

1. Belt, T. G. *et al.* MRI of Wilms' tumor: promise as the primary imaging method. *AJR Am. J. Roentgenol.* **146**, 955–961 (1986).
2. Kangarloo, H., Dietrich, R. B., Ehrlich, R. M., Boechat, M. I. & Feig, S. A. Magnetic resonance imaging of wilms tumor. *Urology* **28**, 203–207 (1986).
3. Gylys-Morin, V., Hoffer, F. A., Kozakewich, H. & Shamberger, R. C. Wilms tumor and nephroblastomatosis: imaging characteristics at gadolinium-enhanced MR imaging. *Radiology* **188**, 517–521 (1993).
4. Koh, D.-M., Takahara, T., Imai, Y. & Collins, D. J. Practical aspects of assessing tumors using clinical diffusion-weighted imaging in the body. *Magn. Reson. Med. Sci. MRMS Off. J. Jpn. Soc. Magn. Reson. Med.* **6**, 211–224 (2007).
5. Humphries, P. D., Sebire, N. J., Siegel, M. J. & Olsen, Ø. E. Tumors in Pediatric Patients at Diffusion-weighted MR Imaging: Apparent Diffusion Coefficient and Tumor Cellularity. *Radiology* **245**, 848–854 (2007).
6. Gawande, R. S., Gonzalez, G., Messing, S., Khurana, A. & Daldrup-Link, H. E. Role of diffusion-weighted imaging in differentiating benign and malignant pediatric abdominal tumors. *Pediatr. Radiol.* **43**, 836–845 (2013).
7. Kim, S. & Chung, D. H. Pediatric solid malignancies: neuroblastoma and Wilms' tumor. *Surg. Clin. North Am.* **86**, 469–487, xi (2006).
8. Aslan, M. *et al.* Diffusion-weighted MRI for differentiating Wilms tumor from neuroblastoma. *Diagn. Interv. Radiol.* **23**, 403–406 (2017).
9. McDonald, K., Sebire, N. J., Anderson, J. & Olsen, O. E. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy-feasibility study. *Pediatr. Radiol.* **41**, 99–106 (2011).
10. Olsen, Ø. E. *et al.* Changes in computed tomography features following preoperative chemotherapy for nephroblastoma: relation to histopathological classification. *Eur. Radiol.* **14**, 990–994 (2004).
11. Gow, K. W. *et al.* Local staging of Wilms' tumor--computerized tomography correlation with histological findings. *J. Pediatr. Surg.* **35**, 677–679 (2000).
12. Littooij, A. S., Humphries, P. D. & Olsen, Ø. E. Intra- and interobserver variability of whole-tumour apparent diffusion coefficient measurements in nephroblastoma: a pilot study. *Pediatr. Radiol.* **45**, 1651–1660 (2015).

13. Littooij, A. S., Sebire, N. J. & Olsen, Ø. E. Whole-tumor apparent diffusion coefficient measurements in nephroblastoma: Can it identify blastemal predominance? *J. Magn. Reson. Imaging* **45**, 1316–1324 (2017).
14. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
15. Weirich, A. *et al.* Clinical impact of histologic subtypes in localized non-anaplastic nephroblastoma treated according to the trial and study SIOP-9/GPOH. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **12**, 311–319 (2001).
16. Beckwith, J. B., Zuppan, C. E., Browning, N. G., Moksness, J. & Breslow, N. E. Histological analysis of aggressiveness and responsiveness in Wilms' tumor. *Med. Pediatr. Oncol.* **27**, 422–428 (1996).
17. Gutjahr, P. *et al.* Klinik, Therapie und Prognose bei 373 Kindern mit Wilms-Tumoren - Ergebnisse der bundesweiten Studie 1980-1988. *Aktuelle Urol.* **21**, 132–141 (1990).
18. Zuppan, C. W., Weeks, D. A., Beckwith, J. B., Luckey, D. W. & Pringle, K. C. The effect of preoperative therapy on the histologic features of wilms' tumor. An analysis of cases from the third national Wilms' tumor study. *Cancer* **68**, 385–394 (1991).
19. Littooij, A. S. *et al.* Apparent diffusion coefficient as it relates to histopathology findings in post-chemotherapy nephroblastoma: a feasibility study. *Pediatr. Radiol.* **47**, 1608–1614 (2017).
20. Meeus, E. M. *et al.* Diffusion-weighted MRI and intravoxel incoherent motion model for diagnosis of pediatric solid abdominal tumors. *J. Magn. Reson. Imaging JMRI* **47**, 1475–1486 (2018).

Chapter 5: Thesis Aims and Objectives

This thesis contains four original research studies which are detailed in Chapters 6 – 9. Overall this thesis aims to investigate the use of diffusion MRI in Wilms' tumours. Currently most research has focused on ADC, derived from a mono-exponential model; thus, this thesis will expand on past research by additionally investigating advanced non-Gaussian models of diffusion. The four models under investigation are: mono-exponential (ADC), IVIM (D, D^*, f), stretched exponential (DDC, α) and kurtosis (D_k, K).

This thesis will first investigate:

1. The reproducibility of the derived parameters from the four diffusion models in healthy renal tissue
2. The goodness of fit of the four diffusion models in Wilms' tumour tissue and the contralateral normal kidney

The thesis will then investigate clinical questions:

1. Can any of the parameters derived from the four diffusion models identify necrotic tissue in Wilms' tumour tissue?
2. Can any of the parameters derived from the four diffusion models separate Wilms' tumour subtypes?

5.1: Aims and Hypotheses

The aims and hypotheses are based on past research which is described in the introductory sections of the following chapters.

5.1.1: Chapter 6: Reproducibility of Diffusion Parameters in Healthy Adult Kidney Tissue

Each parameter from the four diffusion models will be tested to see if they are reproducible when different field strengths are used (1.5T vs. 3T) and when different b values are acquired. This will be tested based on coefficients of variation, parametric and non-parametric tests, and Bland-Altman analysis.

A parameter will be considered to have low variability if the coefficients of variation and levels of bias (as assessed by the mean difference from Bland-Altman analysis) are ≤15%.

Aims and Hypotheses:

- In healthy adult kidneys there will be a difference in reproducibility between the derived parameters from each model.
- For both experiments (field strength and b values) it is expected that the greatest variability will be seen in D^* and f (IVIM) and K (kurtosis).
- The other parameters ADC (mono-exponential), D (IVIM), DDC and α (stretched exponential) and D_k (kurtosis) will have lower variability.

5.1.2: Chapter 7: Comparison of Models of Diffusion in Wilms' Tumour and Contralateral Renal Tissue

Each model will be fit on a voxel-by-voxel basis and the Akaike information criterion (AIC) will be calculated for each model for each voxel. The lowest AIC value will determine which model best describes the diffusion data for each voxel. The model which best describes the majority of the tissue will be classified as the best fit model.

Aims and Hypotheses:

- In contralateral renal tissue IVIM will provide a superior fit to the raw diffusion data compared to the mono-exponential model.
- As there is a lack of research into stretched exponential and kurtosis models in normal renal tissue and a lack of research into all three non-Gaussian models in Wilms' tumour, specific directional hypotheses cannot be made on these models. This study therefore aims to identify which of the models provide the best fit to this data.
- It is expected that the different Wilms' tumour subtypes will favour different models as they have different cellular environments.

5.1.3: Chapter 8: An Alternative Approach to Contrast-Enhanced Imaging: Diffusion Weighted Imaging and T_1weighted Imaging Identifies and Quantifies Necrosis in Wilms' Tumour

Multiple regression analysis will investigate whether a combination of a diffusion parameter and T_1weighted imaging is related to gadolinium enhancement in Wilms' tumour tissue. Models will be compared based on adjusted R^2 values with the highest adjusted R^2 selecting the preferred model.

Aims and Hypotheses:

- There will be a relationship between mean diffusion values and mean fractional enhancement (based on gadolinium contrast) in Wilms' tumours.
- An inverse relationship will be seen in the majority of parameters with an increase in ADC (mono-exponential), D and D^*(IVIM), DDC and α (stretched exponential), D_k and K (kurtosis) being significantly ($p < 0.05$) related to a decrease in mean fractional enhancement, the parameter f x D^* (IVIM) will also be investigated and the same relationship will be seen.
- A positive relationship will be seen with f, with an increase in mean f being significantly ($p < 0.05$) related to an increase in mean fractional enhancement.
- The IVIM parameters are expected to provide the highest adjusted R^2 values.
- The selected model from the regression analysis will be compared to the gadolinium method and will have little bias. A small bias will be defined as below 10% based on Bland-Altman analysis.
- A threshold will be determined which can separate necrotic and viable Wilms' tumour tissue based on diffusion and T_1weighted imaging

5.1.4: Chapter 9: Identification of Wilms' Tumour Subtypes using Non-Gaussian Models of Diffusion

One-way ANOVAs will investigate whether there are significant differences ($p < 0.05$) between the Wilms' tumour subtypes (blastemal, epithelial, stromal, mixed and regressive) based on diffusion values. Post-hoc testing will reveal where these differences

lie. Whole tumour values will be compared to viable tumour values to determine whether limiting the analysis to just the viable tumour tissue improves the subtype separations.

Aims and Hypotheses:

- There will be significant differences between the subtypes based on diffusion values.
- The non-Gaussian models will provide more subtype separations compared to the mono-exponential model
- The blastemal and epithelial subtypes will have lower ADC values compared to the other subtypes.
- The stromal subtype will have the highest ADC values compared to the other subtypes.
- Limiting the analysis to just the viable tumour tissue will provide more significant differences between the subtypes compared to using the whole tumour values.

Chapter 6: Reproducibility of Diffusion Parameters in Healthy Adult Kidney Tissue

This chapter details work from an original research study. It describes the reproducibility of a range of diffusion parameters from mono-exponential and non-Gaussian models in healthy adult renal tissue. This was a prospective study that investigated whether parameters were reproducible when compared between a 1.5T and 3T system and when different b values were acquired. Reproducibility was assessed based on coefficients of variation, parametric and non-parametric tests and Bland-Altman analysis. It found that ADC (mono-exponential), *D* and *f* (IVIM), α (stretched exponential), and D_k (kurtosis) had lower levels of variability (assessed by coefficients of variation and levels of bias - the mean difference from Bland-Altman analysis - ≤15%) compared to *D** (IVIM), *DDC* (stretched exponential) and *K* (kurtosis).

6.1: Background of the Reproducibility of Diffusion Parameters

Diffusion weighted imaging (DWI) has been shown to be a very useful technique in body imaging. ADC has proven to be a clinically important parameter in Wilms' tumour[1,2], and parameters from other non-Gaussians models of diffusion (IVIM[3], stretched exponential[4] and kurtosis[5]) have also demonstrated their utility in a range of body pathologies (full details on these diffusion models and their clinical utility are given in Chapter 3). Given that DWI can potentially produce a range of clinically useful parameters it is important to determine whether these parameters are reproducible, as different institutions and hospitals are likely to acquire data using a range of scanner models, magnetic field strengths, and DWI protocols.

Multi centre studies have investigated the stability of diffusion parameters based on different manufacturers and varying b values in the brain[6]. This demonstrated that both ADC and IVIM parameters are largely reproducible across scanners and field strengths. However, in body imaging the literature is more varied. This may be influenced, in part, by the choice of whether breath-hold, free breathing or respiratory triggered acquisitions are used. Table 6.1 summarises some studies which have investigated reproducibility of diffusion parameters across the body. Low variability is defined as those with values ≤15%.

Table 6.1. A summary of research studies which have investigated reproducibility of diffusion parameters. In this table 'Consistent Acquisition' means there were no differences in b values, field strength or other acquisition methods between scans.

Study Author	Model /Parameter	Organ	Methodology	Main Findings
Braithwaite et al. 2009[7]	ADC	Healthy abdominal tissue: Liver, Spleen Pancreas	Consistent Acquisition	No significant differences in ADC between scans
Thoeny et al. 2005[8]	ADC	Healthy Kidney Tissue	Consistent Acquisition	No significant differences in ADC between scans
Cutajar et al. 2011[9]	ADC	Healthy Kidney Tissue	Consistent Acquisition	No significant differences in ADC between scans and low variability (<15%)
Jakab et al. 2017[10]	IVIM	Healthy Foetal MRI (Kidneys, Liver, Lungs)	Consistent Acquisition	*D** and *f* had greater variability (14 – 25%), *D* was more reproducible (12 – 14%)
Sun et al. 2017[11]	IVIM	Rectal Cancer and Prostate Tissue	Consistent Acquisition	*D** and *f* showed higher variability (126 – 197%) than *D* (24.5%) in rectal cancer, in prostate tissue all IVIM parameters were reproducible (Bland-Altman confidence intervals: <5%)
Reischauer et al. 2017[12]	IVIM & Stretched Exponential	Prostate Cancer Bone Metastases	Consistent Acquisition	*D** and *f* had greater variability (20 – 43%), *D, DDC* and α were all reproducible (<10%)
Winfield et al. 2015[13]	Stretched Exponential	Ovarian Cancer	Consistent Acquisition	*DDC* and α were both reproducible (4.3 – 7%)

Jerome et al. 2017[14]	Mono-exponential, IVIM, Stretched Exponential, & Kurtosis	Range of paediatric tumours, (cranial and extra-cranial)	Consistent Acquisition	Majority of parameters showed good reproducibility (ADC, *D*, *DDC*, α, D_k < 10%). *D*, f* and *K* had the highest variability (>30%)
Pentang et al. 2014[15]	Kurtosis (D_k)	Healthy Kidney Tissue	Consistent Acquisition	D_k no significant differences, however this was based on one volunteer
Kwee et al. 2008[16]	ADC	Healthy Liver Tissue	Breathing Technique Alteration	Breath-hold and free breathing methods provided consistent ADC values. Respiratory triggering gave very variable ADC values
Choi et al. 2013[17]	ADC	Focal Liver Lesions	Breathing Technique Alteration	Breath-hold and free breathing methods provided consistent ADC values
Ye et al. 2014[18]	ADC	Heathy Pancreatic Tissue	Field Strength Alteration	ADC values were higher when acquired on 1.5T scanners compared to 3T
Barral et al. 2013[19]	ADC	Heathy Pancreatic Tissue	Field Strength Alteration	No significant differences in ADC between 1.5T and 3T
Dale et al. 2010[20]	ADC	Healthy Spleen, Liver and Pancreatic Tissue	Field Strength and b value Alteration	No significant differences in ADC were found when field strength was altered. When b value ranges were focused on lower b value weightings, ADC values increased. When the maximum b value was kept consistent ADC did not differ despite the inclusion of different lower b value weightings

Cui et al. 2015[21]	IVIM	Healthy liver tissue and Chronic Liver Disease	Field Strength Alteration	D^* and f had greater variability (22 – 47%), D was more reproducible (10.9%)
Barbieri et al. 2015[22]	IVIM	Heathy Upper Abdominal Organs	Field Strength Alteration	Greatest variability seen in D^* and f (8.7 – 36.7%) in renal cortex and medulla, D was reproducible (<7%)

6.1.1: Diffusion Parameters Reproducibility: Consistent Acquisition

Variations in the acquisition protocol may alter the measured diffusion parameters, as illustrated in Table 6.1; however, it is also important to establish whether these parameters are reproducible when the acquisition protocol is kept constant, which may be more indicative of reproducibility within a single-centre setting.

The mono-exponential model, which produces ADC, appears to be stable when acquisition is unchanged. For example, when the same DWI sequence was repeated five times in a single session using a 3T scanner, ADC did not significantly differ within a given healthy subject between each acquisition when obtained in the liver, spleen and pancreas[7]. There were also no significant changes in ADC values when volunteers returned 147 days later. This has also been demonstrated in the kidneys; 5 healthy volunteers, underwent the same DWI acquisition 6 months apart on a 1.5T scanner, and it was shown that ADC values did not significantly differ between these sessions based on student t-tests[8]. Furthermore, another study on healthy kidneys also demonstrated small within subjects coefficients of variation (CV) for ADC: <10% when subjects were scanned repeatedly on the same day, and when they returned a month later (CV: <15%)[9].

For the non-Gaussian models of diffusion there appears to be greater variability in diffusion parameters even when acquisition parameters do not change. Reproducibility of IVIM parameters have been tested in foetal MRI[10], where a DWI sequence was repeated within a single session in utero. The acquisition protocol consisted of 16 b values (maximum b value = 900 s/mm^2) and the IVIM model was fit using a 2-step method; fitting for D at b values above 250 s/mm^2 and then simultaneously fitting for f and D^* with a fixed D. The kidneys, liver and lungs were selected as volumes of interest, and IVIM values in each organ were compared between the two scans. Across the three organs D^* and f had higher levels of variability (14 – 36%) compared to D (13 – 17%) which remained more stable between the scans. It was suggested that the large range and well-spaced high b value weightings improved the fitting of D, resulting in better repeatability. The higher variability of D^* and f was explained due to the limitations of hardware to achieve a useful range of low b values, furthermore there was difficulty in identifying organs in utero which may have also contributed to the variability in the parameters. It should also be noted that foetal MRI presents unique challenges, and the reproducibility values obtained in this study may be elevated as a result.

Further research has supported the finding that D has higher reproducibility than D^* and f. In a clinical population of patients with rectal cancer[11] it was shown that when

DWI was repeated within a single session, *D* remained more stable (24.5%) compared to *D** and *f*, which had very high levels of variability (197.4% and 126.3%, respectively), and thus caution should be taken when using these measurements. It was suggested that the motion in the rectum could be a potential reason for this region having low repeatability. This study also looked at prostate tissue and found that in this organ all IVIM parameters had better repeatability, with no significant differences between scans, and Bland-Altman analysis also revealed little variability in terms of bias and 95% CI: *D:* bias = 0.8%, (CI: -12 to +13.7), *D**: bias = -4.7%, (CI: -92.4 to +83), and *f* : bias = -3.9%, (-49.1 to +41.4).

The reproducibility of the stretched exponential model has also been investigated in body imaging. Patients with prostate cancer bone metastases were analysed pre-treatment and it was shown that both *DDC* and α were highly reproducible, with coefficients of variation (CV): <10%. The same was true for ADC, and *D* (IVIM)[12]. The authors also highlighted that both *f* and *D** were far less reproducible (CVs: 20%, 43% respectively), which is in line with the previously reported findings[11,23]. Furthermore, *DDC* and α have also been reported to have very low CV (4.3% and 7% respectively) in pre-treatment patients with primary ovarian cancer[13].

Research into the reproducibility of the kurtosis model is limited, particularly in body imaging. However, a study of children (ages 6 to 15 years old) with solid tumours investigated the repeatability of mono-exponential, IVIM, stretched exponential and kurtosis models; this study included both cranial and extracranial tumours[14]. Comparing parameters from two scans (24 hours apart) on a 1.5T scanner it was shown that values of ADC, *D* (IVIM), *DDC* (stretched exponential), α (stretched exponential), and D_k (kurtosis) were all reproducible with CVs below 10%, however *K* (kurtosis), *f* (IVIM), and *D** (IVIM) were far less reproducible with CVs >30%. The selection of b values (b = 0, 50, 100, 300, 600, 1,000 s/mm^2), may have had a role in the variability of these parameters. A range of tissues were analysed (including brain tissue), thus higher b values may have been needed to capture the effects of the kurtosis model. Additionally, there was a limited number of low b values and this may have also impacted the reproducibility of *D** and *f* which are influenced by lower b values. The scan time (6 minutes) was found to be tolerable for these patients and the inclusion of more b values would have lengthened this, which may have caused problems for this paediatric population.

Repeatability of D_k, from the kurtosis model, has also been demonstrated in the kidneys[15]; however this was based on one healthy volunteer who was scanned twice, 14 days apart, using the same acquisition protocol. Mean values of D_k from within the kidneys were compared between the two scans using student's t-tests and no significant

differences were found. However, as this result is based on very limited data further testing of D_k in the kidneys is needed to draw any reliable conclusions.

6.1.2: Diffusion Parameters Reproducibility: Acquisition Alterations – Breathing Techniques

When acquiring DWI data different methods can be used to account for respiratory motion, including free breathing, breath-hold or respiratory triggering. Breath-hold techniques have much shorter acquisition times and as such the signal-to-noise ratio (SNR) is generally lower compared to other methods[24]. It has been suggested, however, that using breath-hold techniques, as opposed to free breathing, may provide more stable ADC values of small lesions as there would be fewer partial volume artefacts related to respiratory motion[25]. However, it has been shown that ADC values in the liver were not significantly different when obtained via breath-hold or free breathing methods; this has been shown in both 1.5T[16] and 3T[17] systems.

Respiratory triggering methods require longer scan times, with increased SNR, but results have shown that ADC values were more variable and less reproducible than free breathing and breath-hold methods, when measured in the liver[16]. The higher ADC values found when using respiratory triggering, in this particular study, may have been due to the higher blood flow rate at the end of expiration, and the authors suggested that DWI signal should be evaluated at different phases of the respiratory cycle[16].

Overall it has been suggested that as free breathing methods are easier to acquire, more time efficient than respiratory triggering, and are reproducible, thus it may be the most appropriate acquisition technique[16,17].

6.1.3: Diffusion Parameters Reproducibility: Acquisition Alterations – Field Strength

Different research and clinical centres are likely to have scanners with varying field strengths. It is thus important to determine whether this will impact the measured diffusion parameters. The impact of field strength on the reproducibility of ADC in the body has been investigated. In a prospective healthy volunteer study of the pancreas, it was shown that ADC values were significantly higher in a 1.5T GE scanner compared to a 3T Philips scanner, when subjects were scanned consecutively using the same b values[18].

This difference also remained when ADC values from a 1.5T Siemens scanner were compared to the 3T Philips scanner. Previously it has been suggested that MRI at 3T has larger magnetic field inhomogeneities and greater susceptibility artefacts[26], and the this may have contributed to the differences seen in pancreatic ADC values between the 1.5T and 3T[18]. However, importantly, the differences in ADC may also have been due to the differences in vendors and the different echo times used, as the sequences were not identical. Thus, it is difficult to isolate the impact of field strength on these measurements. Conversely, another study showed that ADC values of the pancreas did not significantly differ when obtained from 1.5T and 3T systems, in this instance the vendor remained consistent[19]. Furthermore, another study also demonstrated no significant differences in ADC between 1.5T and 3T (both Siemens scanners) in the spleen[20].

Stability of IVIM parameters in the body have also been tested between field strengths. It was shown that when *D*, *D**, and *f* were measured in liver tissue on a 1.5T and also on a 3T (average of 9 days apart) the mean CV (comparing 1.5T values to 3T values) was much greater in *f* (22.6%) and *D** (46.8%) compared to *D* (10.9%)[21]. However, it should be noted that the MR systems were from different vendors; Siemens 1.5T and GE 3T, and this study combined healthy volunteers (n = 2) and patients with chronic liver disease (n = 17). Thus, these factors may have also affected the results. Furthermore, the high variability in *D** was likely to be due to fitting errors, and as mentioned previously[10,11] this parameter has high variability when repeated on the same system, thus it is unsurprising that it differs greatly when measured across different field strengths.

A recent study into the effects of field strength and MR vendor on measured IVIM parameters in upper abdominal organs also provided similar findings[22]. A variety of vendors of 1.5T and 3T systems were included. It was shown that in the renal cortex significant differences were found in *D*, *D** and *f*, between 1.5T and 3T systems, and in the renal medulla significant differences in *D* and *f*. Between-subject and within-subject variation was also investigated, combining all MR systems. Variations in *D* were smaller (5.5-6.8%) than *f* (8.7 – 23.4%) and *D** (28.0 – 36.7%) in the renal cortex and medulla. These variations are similar to the previously mentioned coefficients[21].

6.1.4: Diffusion Parameters Reproducibility: Acquisition Alterations – b values

The range of b values used in the DWI acquisition may also influence reproducibility. A comparison study investigated the effect of different b value ranges on

the measured ADC in the spleen, liver and pancreas[20]. Four b value ranges were used: 0/50 s/mm^2, 0/400 s/mm^2, 0/800 s/mm^2, and 0/50/400/800 s/mm^2 and ADC values were measured using each range in the three abdominal organs. ADC significantly decreased between the 0/50 s/mm^2, 0/400 s/mm^2, and 0/800 s/mm^2 ranges in every organ. This was possibly due to the perfusion effects which are more prominent at lower b value weightings[3]. It was suggested that b value ranges which focused on lower b values would artificially elevate the ADC due to the perfusion effects, and thus ranges which included more acquisitions at higher b values would be less influenced by this[20]. When comparing the 0/800 s/mm^2 and 0/50/400/800 s/mm^2 range there were no significant differences in ADC in the pancreas and the spleen; there was, however, a small difference in the liver. Thus, using different b values may affect the measured ADC, however if the maximum b value is constant, then these differences may be reduced.

6.2: Aims and Hypotheses

It is clear that different diffusion parameters result in varying levels of reproducibility. While research into ADC variability (including changes with field strength[18–20], and b values[20]) has been well studied, the literature on IVIM, stretched exponential and kurtosis is more sparse when the focus is on body imaging.

As previously mentioned DWI is often obtained in patients with Wilms' tumour, and ADC has been shown to provide clinically useful information regarding treatment response and subtype idenfticaion[1,2]. Non-Gaussian models of diffusion have not been greatly investigated in Wilms' tumour, but these parameters may also provide additional information regarding tissue status. Before diffusion models are further investigated in Wilms' tumour it is important to establish how reproducible these parameters are in healthy renal tissue.

The present study sought to determine whether diffusion parameters were reproducible in healthy adult kidney tissue. The study focused on two areas of reproducibility: field strength and b value selection. Diffusion parameters from the four previously mentioned diffusion models (mono-exponential, IVIM, stretched exponential and kurtosis) were assessed to determine which remained stable when calculated based on different scanners of different field strengths (1.5T vs. 3T), and also when different b values were used within the same scanner (0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2 vs. 0, 50, 100, 250, 500, 750, 1000 s/mm^2 on the 3T system.)

Low variability was defined as coefficients of variation and levels of bias (as assessed by the mean difference from Bland-Altman analysis) being ≤15%. It was expected that in healthy adult kidneys there would be a difference in reproducibility between the derived parameters from each model. For both experiments (field strength and b values) it was expected that the greatest variability would be seen in D^* and f (IVIM) and K (kurtosis) based on previous research[11,12,14,21,23]. The other parameters ADC (mono-exponential), D (IVIM), DDC and α (stretched exponential) and D_k (kurtosis) were expected to have lower variability.

6.3: Methods

6.3.1: Study Population

Data was prospectively acquired for this single centre reproducibility study (May 2017 – January 2018). All volunteers attended Great Ormond Street Hospital (London, UK) on two separate occasions for each MRI scan. Volunteers had no known medical conditions and where possible each volunteer came at a similar time of day to their previous scan, having eaten and drank approximately the same quantity as before.

6.3.2: MRI

Imaging was performed on a 1.5T Siemens Magnetom Avanto and on a 3T Siemens Magnetom Prisma both located at Great Ormond Street Hospital. DWI was obtained on each scanner. The 1.5T DWI protocol was as follows: slice thickness: 6mm, TR/TE: 2700ms/89ms, voxel size: 1.4 x 1.4 x 6mm, and 8 b values in 3 orthogonal directions (0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2), in an axial plane. Two DWI protocols were obtained on the 3T, each using identical parameters, except for the number of b values. The 3T protocol was as follows: slice thickness: 5mm, TR/TE: 5000ms/58ms, voxel size: 0.9 x 0.9 x 5mm, and either 7 (0, 50, 100, 250, 500, 750, 1000 s/mm^2) or 8 b values in 3 orthogonal directions (the 8 b values matched that of the 1.5T protocol). Each b value had 9 averages and trace images were used for analysis. The acquisition protocols matched the clinical acquisitions used at Great Ormond Street Hospital.

6.3.3: Data Processing and Regions of Interest

Diffusion data was processed using the trace images on in-house routines designed in Matlab 2015b (MathWorks Inc., Natick, MA, USA) on a voxel-by-voxel basis using four different models of diffusion: mono-exponential, IVIM[3] (Equation: 3.4), stretched-exponential[4] (Equation: 3.5), and kurtosis[5] (Equation: 3.6). The mono-exponential model was fit using the 'polyfit' and 'polyval' functions in Matlab, which is a standard method of fitting a polynomial.

For the non-Gaussian models, fitting was performed using the Levenberg-Marquardt nonlinear least squares algorithm, across all b values (except for the IVIM model). In each case S_0 was defined at the signal at b=0. For the IVIM model, firstly, a linear fit of ln(S/S0) against b was calculated at high b values (200 – 1,000 s/mm^2) to determine the value of D. Following this D^* and f were fit simultaneously (with a fixed D). D^* had no constraints on upper boundaries, and f was constrained between 0 – 1. For the stretched exponential model DDC had no upper boundary conditions, and α was constrained between 0 – 1. For kurtosis neither D_k nor K were constrained by upper boundaries, and K had a lower bound of 0.

Regions of Interest (ROIs) were generated using Mango Software (Research Imaging Institute, UTHSCSA). ROIs were drawn on b0 (non-diffusion weighted) images around each kidney for each volunteer. Areas of high flow such as the areas which surround the renal pelvis were excluded, an example can be seen in Figure 6.1. All reproducibility analysis was based on data within the ROIs.

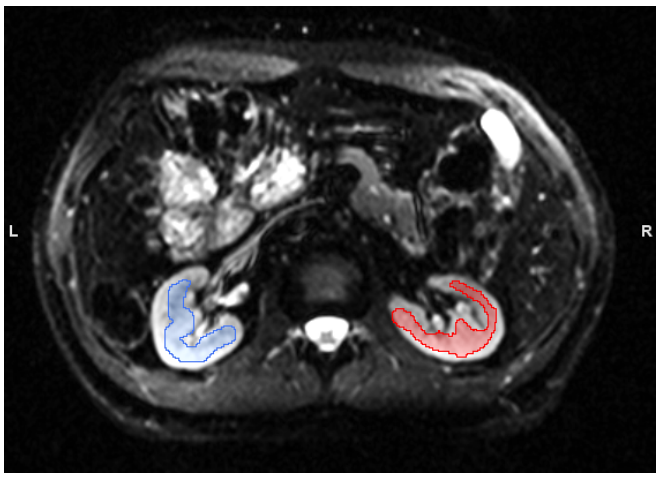

Figure 6.1 A central axial slice of a b0 image from the 3T scanner displaying the abdomen at the level of the kidney from a representative healthy control volunteer (male, 28.26 years). ROIs are drawn around the left (blue) and right (red) kidneys avoiding areas of high flow such as the areas which surround the renal pelvis. In this diagram the right side of the image (white R) is the volunteers' right side.

6.3.4: Analysis: SNR

Signal-to-noise ratio (SNR) was calculated for each volunteer on b0 images and then averaged across volunteers to give an SNR value for the 1.5T and 3T systems. ROIs were placed outside of the abdomen on the b0 images to represent the noise and the kidney ROIs represented the signal. SNR was calculated according to Equation 6.1:

$$\text{SNR} = \frac{0.66^{*}Mean\ Signal}{Mean\ Noise\ Standard\ Deviation} \quad [6.1]$$

As magnitude images were used, the value of 0.66 in Equation 6.1 was used as the Rayleigh distribution correction factor.

6.3.5: Analysis: Reproducibility of Diffusion Parameters

Eight diffusion parameters were generated from the four diffusion models: ADC (mono-exponential), D (IVIM), D^{*} (IVIM), f (IVIM), DDC (stretched exponential), α (stretched exponential), D_k (Kurtosis) and K (Kurtosis). These parameters were generated on voxel-wise basis for each participant; one set of parameters from the 1.5T DWI data, and two sets of parameters from the 3T DWI data: calculated based on the 7 b values and

the 8 b values. In order to compare reproducibility of diffusion parameters between scanners, parameters generated from the 1.5T were compared to the parameters generated from the 3T with matching b values (8 b values). To investigate the effects on stability of diffusion parameters based on different b values, diffusion parameters generated from the 8 b value 3T dataset and 7 b value 3T dataset were compared.

Descriptive statistics of the parameter of interest (means, standard deviations, medians and interquartile ranges), using voxels within the kidney ROIs, were calculated for each participant. The reproducibility of parameters was assessed based on coefficients of variation (CV), parametric and non-parametric statistical tests, and Bland-Altman plots.

6.3.6: Analysis: Reproducibility of Diffusion Parameters – Coefficients of Variation

For each volunteer, mean values for a given parameter (such as ADC) were calculated across all voxels within the kidney ROIs, at each field strength (1.5T and 3T). The mean (μ) and standard deviation (σ) of these two values were used to calculate the coefficient of variation (CV), as follows:

$$CV = \left(\frac{\sigma}{\mu}\right) * 100 \quad [6.2]$$

This process was repeated for each volunteer. The overall CV for each parameter, based on field strength, was then generated by taking the mean of all of the volunteers' CV values.

A similar process was used to calculate the CV across 3T acquisitions acquired using different b values (7 b value data vs. 8 b value data). For each volunteer, the mean fitted value of a given parameter (across all kidney voxels) was calculated using data from both the 7 b value and 8 b value acquisitions. The mean and standard deviation of these were used in Equation 6.2 to calculate a subject-wise CV for that parameter, which was then averaged across all subjects.

6.3.7: Analysis: Reproducibility of Diffusion Parameters – Parametric and Non-Parametric Statistical Tests

To determine whether parameters differed with statistical significance ($p < 0.05$) between scanners or when based on different b values, paired parametric and non-parametric tests were used.

Mean values of one diffusion parameter from each subject from the 1.5T data was used as one sample, and mean values of the same diffusion parameter from the 3T data was used as the second sample. The Lilliefors test was used to determine whether the difference between the two samples was normally distributed. If this was not normally distributed ($p < 0.05$) then a Wilcoxon Signed Rank test was used, if it was normally distributed then a paired Students' t-test was used. The same methods were used to compare parameters generated from 7 b values and 8 b values.

6.3.8: Analysis: Reproducibility of Diffusion Parameters – Bland-Altman Plots

Bland-Altman plots were also used to determine the reproducibility coefficient (1.96 x standard deviation of difference in measured values) and level of bias (mean difference) between the parameters obtained from the 1.5T and 3T scanners, and also between those based on 7 b values and 8 b values.

6.4: Results

6.4.1: Study Population

A total of 5 healthy volunteers were included in the final cohort. Number of cases excluded with reasons are provided in a flowchart (Figure 6.2). Volunteer characteristics can be seen in Table 6.2. The average age of participants on their 1.5T scan was 33.01 years (*SD*: 12.64), and on their 3T scan was 33.10 (*SD*: 12.65).

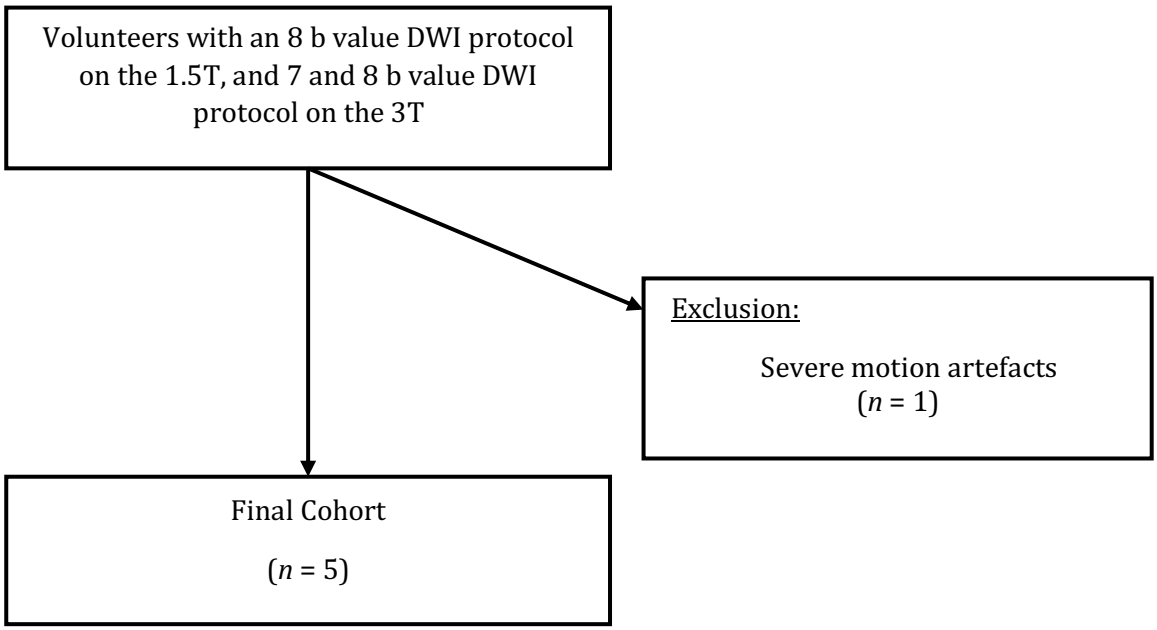

Figure 6.2. Flowchart highlighting inclusions and exclusions of volunteers. DWI = diffusion weighted imaging

Table 6.2. Volunteer characteristics

Volunteer ID	Gender	Age at 1.5T Scan (Years)	Age at 3T Scan (Years)	Days Between Scans
01	Male	28.16	28.26	38
06	Male	55.50	55.60	38
07	Female	25.16	25.27	38
09	Female	27.63	27.74	38
10	Male	28.60	28.62	10

6.4.2: SNR

The average signal (defined as the kidney ROIs on b0 images) across the five volunteers for 1.5T = 188.66 (*SD* = 2.82) and 3T = 216.43 (*SD* = 13.35). The mean noise (defined as an ROI outside of the abdomen on b0 images) across the five volunteers for 1.5T = 2.37(*SD* = 0.27) and 3T = 4.64 (*SD* = 0.28). Equation 6.1 was used to calculate the SNR for 1.5T and 3T. The above average signal values were used for the 'Mean Signal', and the *SD* for the noise was calculated for each volunteer and then the averaged across the five volunteers to provide the measure for 'Mean noise SD'. Using equation 6.1, the SNR for the 1.5T = 82.9, and SNR for the 3T = 57.1.

6.4.3: Descriptive Statistics

Descriptive statistics were calculated for each participant for each parameter. These descriptive statistics were averaged over the 5 volunteers; group means, medians and interquartile ranges can be seen in Tables 6.3 and 6.4. It can be seen that *DDC* had very high mean values and comparatively small median values. This was due to the fitting methods causing some voxels to have extremely high values, thus causing the mean to have high values. Further details of this are given in Section 6.5.

Table 6.3. Descriptive statistics for each field strength (1.5T or 3T) for eight diffusion parameters. Data is averaged across 5 healthy adult kidneys

Field Strength		*ADC* (x 10^{-3} mm^2/s)	*D* (x 10^{-3} mm^2/s)	*D** (mm^2/s)	*f*	D_k (x 10^{-3} mm^2/s)	*K*	*DDC* (mm^2/s)	α
1.5T	Mean	1.81	1.62	0.11	0.27	3.39	0.76	1.94	0.74
	(Std)	(6.31 x 10^{-5})	(5.35 x 10^{-5})	(0.09)	(0.02)	(1.71 x 10^{-4})	(0.04)	(0.34)	(2.01 x 10^{-2})
3T	Mean	1.86	1.63	0.36	0.33	3.79	0.61	1.85 x 10^{+4}	0.79
	(Std)	(1.58 x 10^{-4})	(3.59 x 10^{-5})	(0.10)	(0.11)	(1.66 x 10^{-3})	(0.10)	(3.2 x 10^{+5})	(3.85 x 10^{-2})
1.5T	Median	1.81	1.63	11.65 x 10^{-3}	0.26	3.16	0.79	2.37 x 10^{-3}	0.75
3T	Median	1.84	1.66	8.72 x 10^{-3}	0.29	3.20	0.70	2.46 x 10^{-3}	0.80
1.5T	Interquartile Range	0.15	0.18	1.3 x 10^{-2}	0.10	0.81	0.10	4.25 x 10^{-4}	0.13
3T	Interquartile Range	0.21	0.18	1.47 x 10^{-2}	0.21	2.09	0.29	1.71 x 10^{-4}	0.21

Note. Eight identical b values are used for each field strength. *ADC* = apparent diffusion coefficient (mono-exponential), *D* = 'slow' diffusion coefficient (IVIM), *D** = 'fast' diffusion coefficient (IVIM), *f* = perfusion fraction (IVIM), D_k= diffusion kurtosis Coefficient (kurtosis), *K* = diffusional kurtosis (kurtosis), *DDC* = distributed diffusion coefficient (stretched exponential), and α = stretching parameter (stretched exponential)

Table 6.4. Descriptive statistics for eight diffusion parameters generated from either 7 or 8 b values. Data is averaged across 5 healthy adult kidneys.

Number of b values		*ADC* (x 10^{-3} mm^2/s)	*D* (x 10^{-3} mm^2/s)	*D** (mm^2/s)	*f*	D_k (x 10^{-3} mm^2/s)	*K*	*DDC* (mm^2/s)	α
7	Mean	1.73	1.52	0.31	0.32	3.49	2.51	4.41 x 10^{+6}	0.74
	(Std)	(2.76 x 10^{-4})	2.55 x 10^{-4})	(0.19)	(0.04)	(3.32 x 10^{-4})	(4.04)	9.72 x 10^{+6}	(5.64 x 10^{-2})
8	Mean	1.86	1.63	0.36	0.33	3.79	0.61	1.85 x 10^{+4}	0.79
	(Std)	(1.58 x 10^{-4})	(3.59 x 10^{-5})	(0.10)	(0.11)	(1.66 x 10^{-3})	(0.10)	(3.2 x 10^{+5})	(3.85 x 10^{-2})
7	Median	1.79	1.58	1.06 x 10^{-2}	0.28	2.97	0.79	2.25 x10^{-3}	0.76
8	Median	1.84	1.66	8.72 x 10^{-3}	0.29	3.20	0.70	2.46 x 10^{-3}	0.80
7	Interquartile Range	0.27	0.34	2.16 x 10^{-2}	0.18	1.27	0.26	7.13 x 10^{-4}	0.25
8	Interquartile Range	0.21	0.18	1.47 x 10^{-2}	0.21	2.09	0.29	1.71 x 10^{-3}	0.21

Note. 7 b values =0, 50, 100, 250, 500, 750, 1000 s/mm^2, 8 b values =0, 50, 100, 150, 200, 250, 500, 1000, both obtained on a 3T scanner. *ADC* = apparent diffusion coefficient (mono-exponential), *D* = 'slow' diffusion coefficient (IVIM), *D** = 'fast' diffusion coefficient (IVIM), *f* = perfusion fraction (IVIM), D_k= diffusion kurtosis Coefficient (kurtosis), *K* = diffusional kurtosis (kurtosis), *DDC* = distributed diffusion coefficient (stretched exponential), and α = stretching parameter (stretched exponential)

6.4.4: Coefficients of Variation

The coefficients of variation (calculated according to Equation 6.2) per parameter and volunteer can be seen in Table 6.5 as a function of field strength, and Table 6.6 as a function of b value. *D** and *DDC* had much higher CVs (78.4% and 108.4%, respectively) compared to the other parameters when field strength was being compared, whereas *DDC* and *K* had high and moderately high CVs (116.4% and 31.82%, respectively) when b values were being compared. *DDC* had near identical CV for each volunteer, this was due to the very high mean values of *DDC*, which were highest on the 3T system, and also the very high *SD* of this parameter. These high values are due to the fitting method and are discussed in Section 6.5

Table 6.5. Coefficients of variation for all fitted diffusion parameters, when generated from 1.5T and 3T systems using matched b-values, based on data from 5 healthy subjects

	Coefficient of Variation (%) Between 1.5T & 3T Per Volunteer							
Volunteer ID	*ADC*	*D*	*D**	*f*	D_k	*K*	*DDC*	α
01	0.83	4.44	59.0	5.28	19.49	31.62	141.41	12.10
06	0.49	1.58	129.17	3.10	1.41	4.83	6.14	1.44
07	2.22	0.88	103.93	0.66	7.41	8.92	119.40	4.15
09	9.99	3.12	56.13	44.58	47.77	25.69	141.42	3.82
10	0.95	0.46	43.74	12.40	3.23	9.63	133.77	3.33
MEAN CV (%)	2.89	2.10	78.40	13.20	15.86	16.14	108.43	4.97

Table 6.6. Coefficients of variation for all fitted diffusion parameters, when generated based on 7 b values and 8 b values acquired at 3T, based on data from 5 healthy subjects

	Coefficient of Variation (%) Between 7 b values & 8 b values Per Volunteer							
Volunteer ID	*ADC*	*D*	*D**	*f*	D_k	*K*	*DDC*	*a*
01	23.17	29.12	13.57	30.88	33.45	127.29	141.19	19.70
06	0.80	0.04	14.91	0.41	0.38	1.61	141.42	0.79
07	0.40	1.67	17.60	2.60	2.72	5.07	141.42	1.31
09	6.39	5.55	28.35	37.87	47.92	24.31	16.53	1.95
10	0.59	3.21	64.92	12.00	3.50	0.83	141.42	1.29
MEAN CV (%)	6.27	7.92	27.87	16.74	17.60	31.82	116.40	5.01

6.4.5: Parametric and Non-Parametric Statistical Tests – Field Strength

For comparisons based on field strength the difference between the two samples for all parameters other than ADC, *f*, and D_k were normally distributed ($p > 0.05$), thus Wilcoxon Signed Rank tests were used for these three parameters; all other parameters were tested using paired t-tests. Ladder plots for field strength comparisons can be seen in Figure 6.3.

All parameters other than *D** and *K* did not differ significantly ($p > 0.05$) between the 1.5T and 3T scanners. *D** was significantly higher on the 3T (*Mean:* 0.36, *SD:* 0.1) compared to the 1.5T (*Mean:* 0.11, *SD:* 0.09), $t(4) = -6.08$, $p = 3.7 \times 10^{-3}$. *K*, however, was significantly higher on the 1.5T (*Mean:* 0.76, *SD:* 0.04) compared to the 3T (*Mean:* 0.61, *SD:* 0.10), $t(4) = 3.29$, $p = 3.01 \times 10^{-2}$.

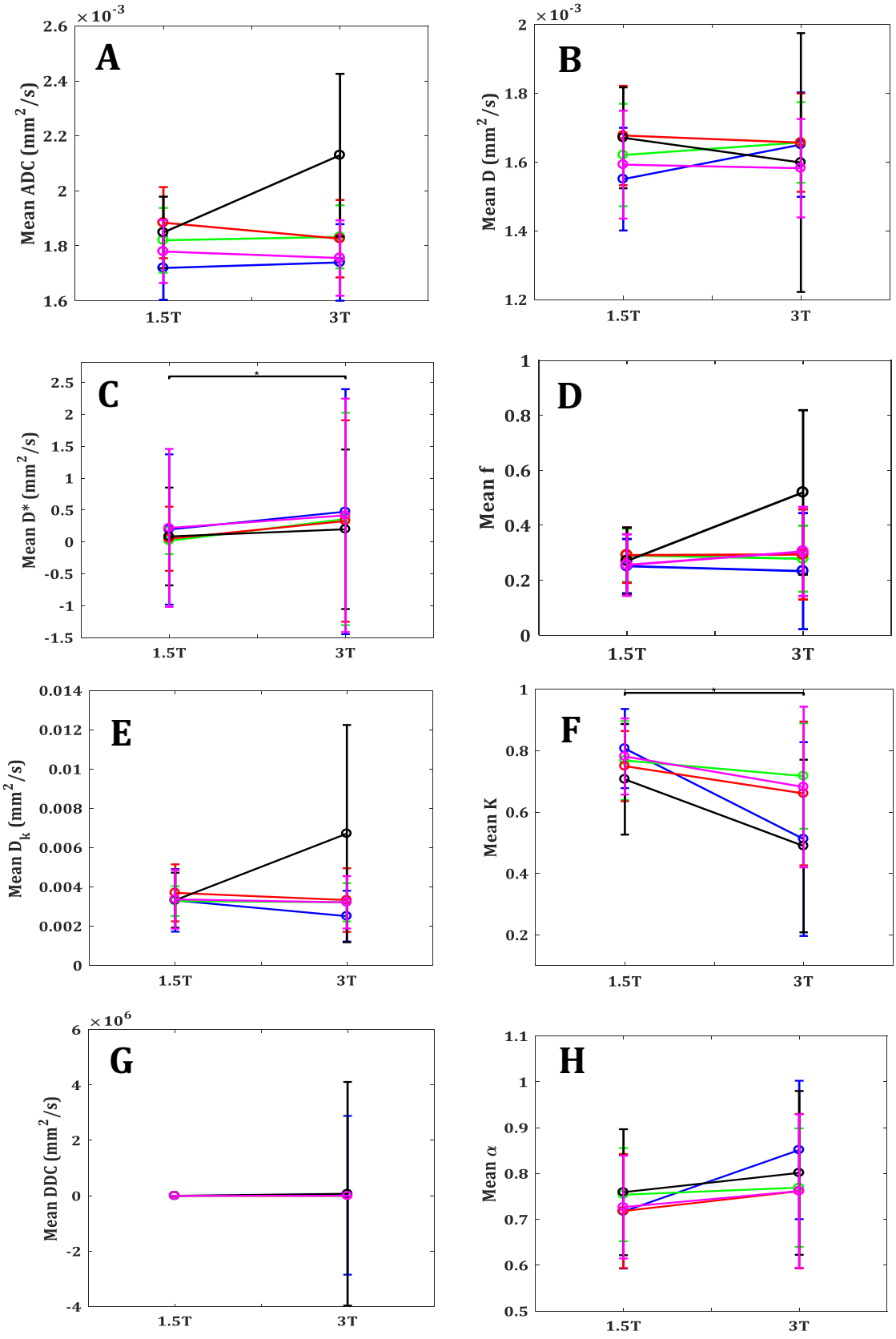

Figure 6.3. Ladder plots for all parameters, comparing mean diffusion parameters on 1.5T and 3T systems. * = significant differences $p < 0.05$. Error bars are 1SD**.** Each colour bar represents one volunteer (blue = 01, green = 06, red = 07, black = 09, pink = 10). **(A):** *ADC* = apparent diffusion coefficient (mono-exponential), **(B):** *D* = 'slow' diffusion coefficient (IVIM), **(C):** *D** = 'fast' diffusion coefficient (IVIM), **(D):** f = perfusion fraction (IVIM), **(E):** D_k = diffusion kurtosis coefficient (kurtosis), **(F):** *K* = diffusional kurtosis (kurtosis), **(G):** *DDC* = distributed diffusion coefficient (stretched exponential), **(H):** α = stretching parameter (stretched exponential)

6.4.6: Parametric and Non-Parametric Statistical Tests – b Values

For comparisons based on b values the difference between the two samples for most parameters (ADC, f, D_k, K, DDC, α) were not normally distributed ($p > 0.05$), thus Wilcoxon Signed Rank tests were used for these parameters, and paired t-tests were used for the remaining parameters. There were no significant differences for any of the parameters when they were compared based on whether they were calculated from 7 or 8 b values. Ladder plots for b value comparisons can be seen in Figure 6.4.

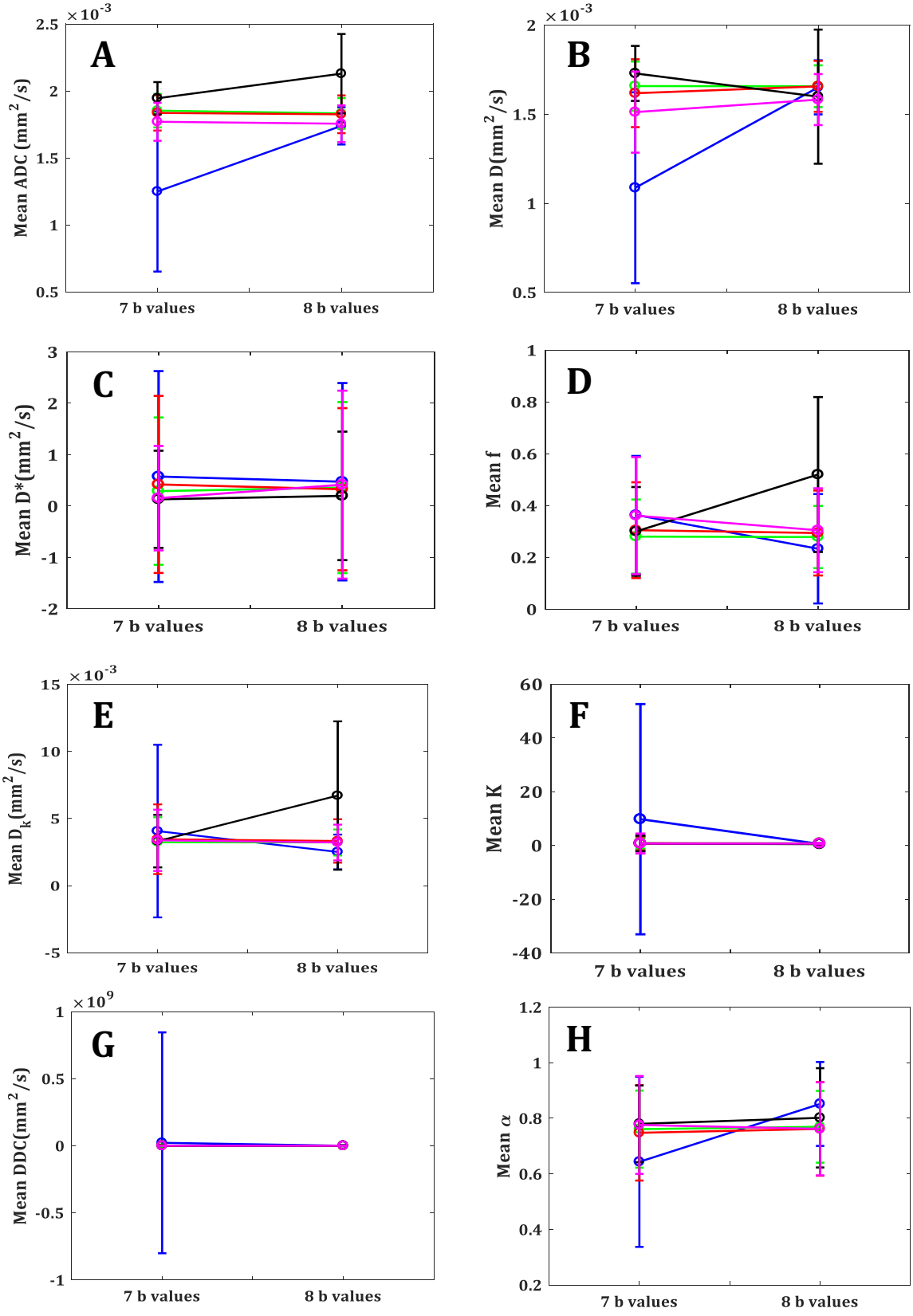

Figure 6.4. Ladder plots for all parameters, comparing mean diffusion parameters generated from 7 b values and 8 b values. * = significant differences $p<0.05$. Error bars are 1SD**.** Each colour bar represents one volunteer (blue = 01, green = 06, red = 07, black = 09, pink = 10). **(A):** *ADC* = apparent diffusion coefficient (mono-exponential), **(B):** *D* = 'slow' diffusion coefficient (IVIM), **(C):** *D** = 'fast' diffusion coefficient (IVIM), **(D):** *f* = perfusion fraction (IVIM), **(E):** D_k = diffusion kurtosis coefficient (kurtosis), **(F):** *K* = diffusional kurtosis (kurtosis), **(G):** *DDC* = distributed diffusion coefficient (stretched exponential), **(H):** α = stretching parameter (stretched exponential)

6.4.7: Bland-Altman Plots – Field Strength

Eight Bland-Altman plots demonstrating the level of agreement between mean parameter values calculated on the 1.5T and 3T systems can be seen in Figure 6.5.

For most parameters there was good agreement between the two systems. ADC, *D,* and D_k (Figures 6.3.A, 6.3.B, 6.3.E, respectively), all had biases <1% between the scanners (a positive mean difference would indicate values were higher at 3T and a negative mean difference would indicate values were higher at 1.5T). ADC: Mean difference = 4.64×10^{-5}, 95% CI= [-2.18×10^{-4}, 3.11×10^{-4}], *D*: Mean difference = 6.79×10^{-5}, 95% CI= [-1.2×10^{-4}, 1.34×10^{-4}], D_k: Mean difference = -4.01×10^{-4}, 95% CI= [-2.92×10^{-3}, 3.72×10^{-3}].

There were also very small bias (5.41%) in *f* and α (Figure 6.3.D & 6.3.H), *f*: Mean difference = 0.0541, 95% CI= [-0.17, 0.27]. α*:* Mean difference = 0.0541, 95% CI= [-3.65×10^{-2}, 0.145]. There was a moderate level of bias (15.03%) in *K* (Figure 6.3.F) whereby the 1.5T produced higher values: Mean difference = -0.1503 95% CI= [-0.35, 0.05].

There was a larger bias (24.2%) in *D** values (Figure 6.3.C) where the 3T produced higher values; mean difference = 0.242, 95% CI= [0.07, 0.42]. In *DDC* (Figure 6.3.G) there was no agreement between the two systems, with the 3T producing vastly higher values; Mean difference = 18450.09, 95% CI= [-444345.8, 81245.97].

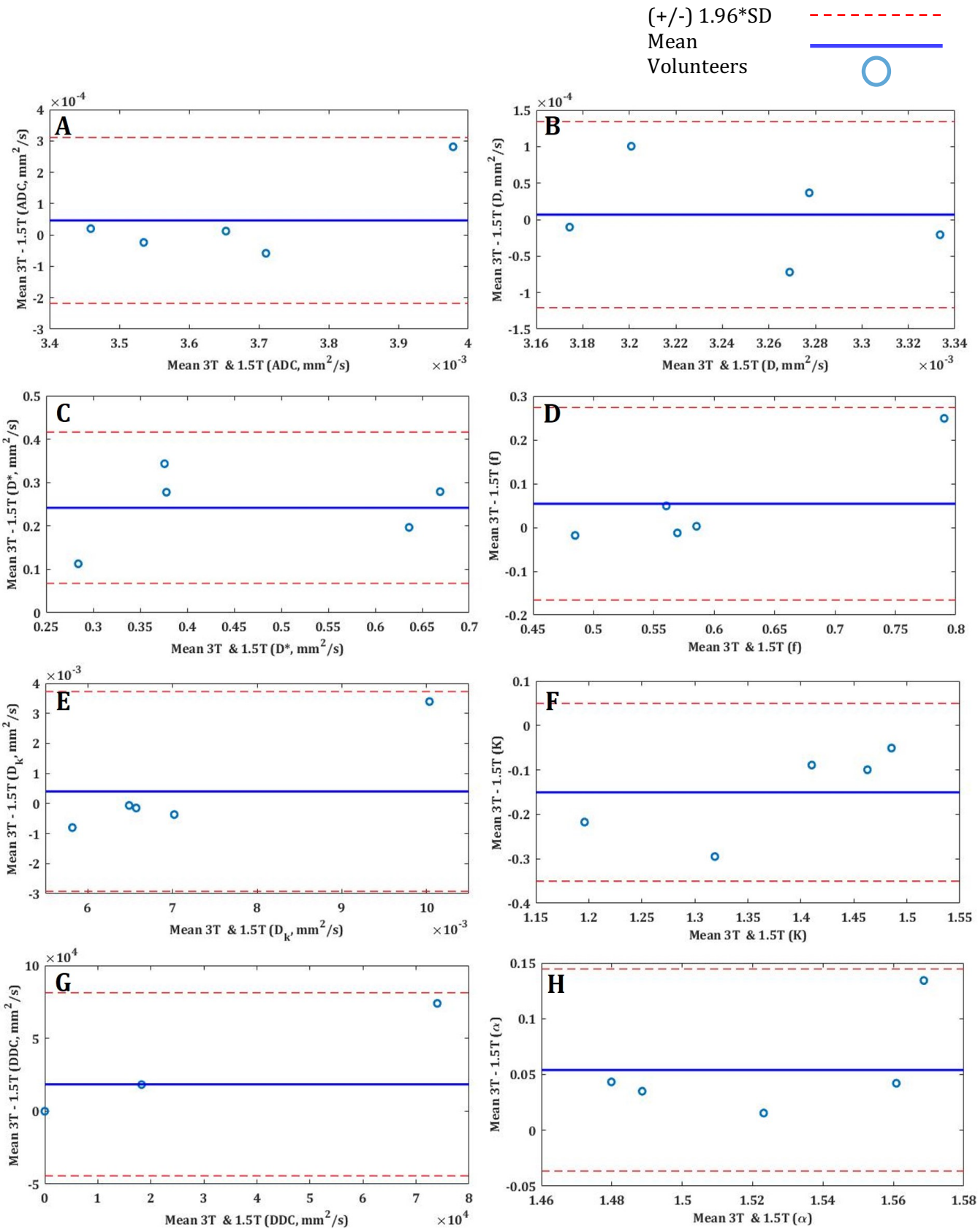

Figure 6.5. Bland-Altman analyses showing the level of agreement for each diffusion parameter when calculated on 1.5T and 3T systems. **(A):** *ADC* = apparent diffusion coefficient (mono-exponential), **(B):** *D* = 'slow' diffusion coefficient (IVIM), **(C):** *D** = 'fast' diffusion coefficient (IVIM), **(D):** f = perfusion fraction (IVIM), **(E):** D_k = diffusion kurtosis coefficient (kurtosis), **(F):** *K* = diffusional kurtosis (kurtosis), **(G):** *DDC* = distributed diffusion coefficient (stretched exponential), **(H):** α = stretching parameter (stretched exponential)

6.4.8: Bland-Altman Plots – b values

Eight Bland-Altman plots demonstrating the level of agreement between mean parameter values calculated based on 7 b values and 8 b values can be seen in Figure 6.6.

For most parameters there was good agreement between the two b value ranges. ADC, *D*, *f* and D_k (Figures 6.4.A, 6.4.B, 6.4.D, 6.4.E, respectively), all had biases <1%; (a positive mean difference would indicate values were higher when calculated based on 8 b values and a negative mean difference would indicate values were higher when calculated based on 7 b values). ADC: Mean difference = 1.26×10^{-4}, 95% CI= [-3.1×10^{-4}, 5.59×10^{-4}], *D*: Mean difference = 1.08×10^{-4}, 95% CI= [-4.1×10^{-4}, 6.29×10^{-4}], *f*: Mean difference = -4.05×10^{-3}, 95% CI= [-0.25, 0.26]. D_k: Mean difference = 3.06×10^{-4}, 95% CI= [-3.3×10^{-3}, 3.91×10^{-3}].

There was a small bias in *D** (4.05%) and α (4.77%) (Figures 6.4.C and 6.4.H, respectively), with both parameters being higher when calculated based on 8 b values; *D**: Mean difference = 0.0406, 95% CI= [-0.25, 0.33], α: Mean difference = 0.0477, 95% CI= [-0.13, 0.23].

Both *K* and *DDC* showed no agreement when calculated based on 7 b values compared to 8 b values (Figures 6.4.F and 6.4.G respectively). Values of *K* were far higher when based on 7 b values (Mean difference = -1.90, 95% CI= [-9.23, 6.13]). *DDC* values were also greater when 7 b values were used (Mean difference = -4393023, 95% CI= [$-2.3 \times 10^{+7}$, $1.5 \times 10^{+7}$]).

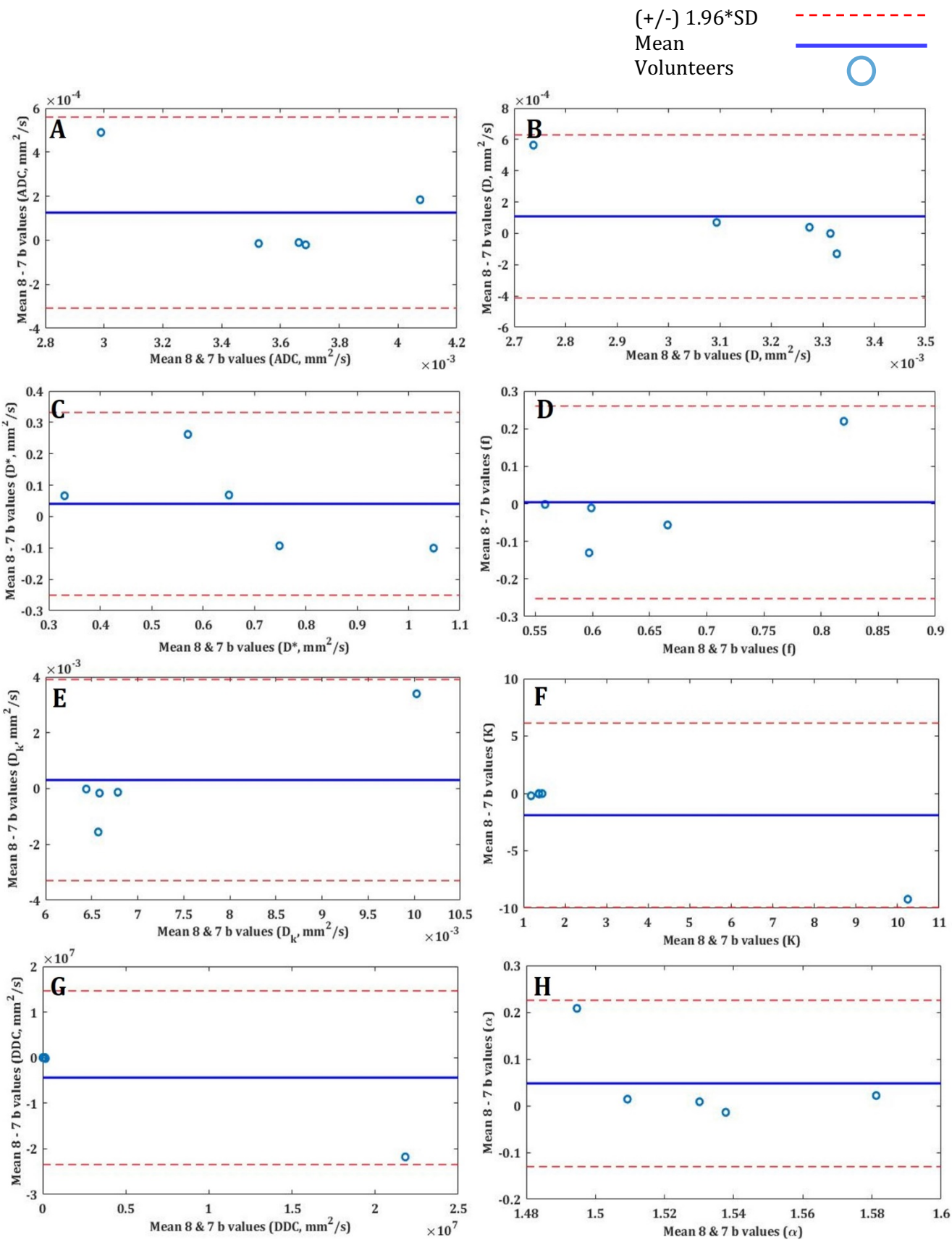

Figure 6.6. Bland-Altman analyses showing the level of agreement for each diffusion parameter when calculated based on 7 and 8 b values. **(A):** *ADC* = apparent diffusion coefficient (mono-exponential), **(B):** *D* = 'slow' diffusion coefficient (IVIM), **(C):** *D** = 'fast' diffusion coefficient (IVIM), **(D):** f = perfusion fraction (IVIM), **(E):** D_k = diffusion kurtosis coefficient (kurtosis), **(F):** *K* = diffusional kurtosis (kurtosis), **(G):** *DDC* = distributed diffusion coefficient (stretched exponential), **(H):** α = stretching parameter (stretched exponential)

6.4.9 – DDC Results

Mean values of *DDC* were very high, this was due to a subset of outlier voxels with extremely high values. The reasoning for this is due to the applied fitting method, which is discussed in Section 6.5. As a way of highlighting the number of voxels with these high values, the percent of voxels within the ROIs which had *DDC* values greater than the diffusivity of free water at body temperature (3×10^{-3} mm^2/s)[27] were calculated for each volunteer for the 1.5T data, and the two b value ranges at 3T. Table 6.7 shows these results, along with the maximum *DDC* values.

Table 6.7. Percent of voxels within the kidney ROIs which showed *DDC* values above diffusivity of free water at body temperature and maximum *DDC* values within the ROIs

	1.5T		3T (8 b values)		3T (7 b values)	
Volunteer ID	% High Voxels	Maximum DDC Value (mm^2/s)	% High Voxels	Maximum DDC Value (mm^2/s)	% High Voxels	Maximum DDC Value (mm^2/s)
01	8.17	1589.44	5.16	$4.5 \times 10^{+8}$	15.57	$9.84 \times 10^{+10}$
06	6.29	0.05	9.0	2.37	7.57	$1.1 \times 10^{+9}$
07	16.45	0.18	13.71	535.07	15.12	$4.17 \times 10^{+8}$
09	12.15	1.22	55.55	$2.92 \times 10^{+8}$	10.32	$2.55 \times 10^{+8}$
10	8.86	432.07	12.59	43.85	15.59	$1.07 \times 10^{+8}$
Mean	10.38%	404.59	19.20%	$1.48 \times 10^{+8}$	12.83%	$2.01 \times 10^{+10}$

An example of a *DDC* map and fits of two voxels is shown in Figure 6.7. It can be seen that in one voxel the signal steadily declines with increasing b value, however in another voxel there is a very quick decline in signal followed by a decrease in signal attenuation. By applying the stretched exponential model to this particular voxel, the result is a very high *DDC* and low α. This may be due to high inhomogeneity within the voxel and a more complex tissue environment, however without histology this only speculation.

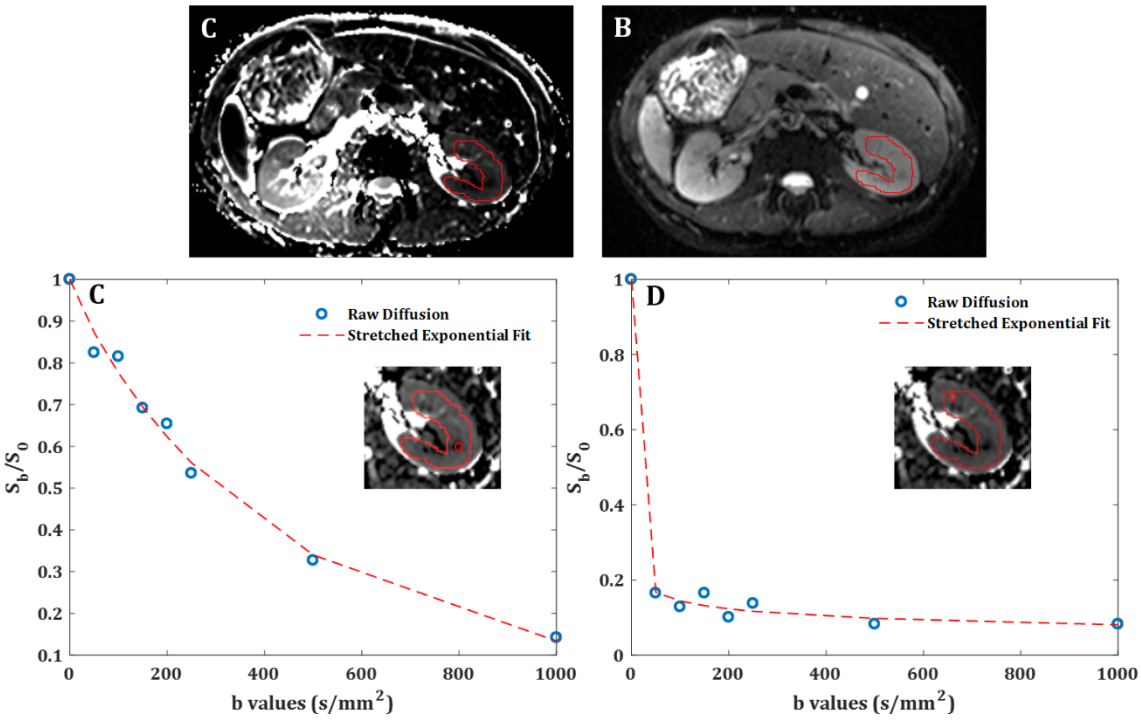

Figure 6.7 A central axial slice of a *DDC* map **(A)** and the corresponding b0 image **(B)** from the 3T scanner displaying the abdomen at the level of the kidney from a representative healthy control volunteer (female, 27.74 years at time of scan). An ROI is shown in red surrounding the healthy renal tissue. The raw diffusion data and stretched exponential fit is plotted from two voxels within the ROI **(C)** shows a voxel with a *DDC* = 2.2 x 10^{-3} mm^2/s and α = 0.9. **(D)** shows a voxel with a *DDC* = 3.27 mm^2/s and α = 0.1

6.5: Discussion

This study compared the reproducibility of eight diffusion parameters from four diffusion models, in five pairs of healthy adult kidneys. Parameters were compared based on differences in field strength and different b values. In terms of field strength, it was demonstrated that *DDC* was highly unstable, showing an extremely high coefficient variation (108%) and no agreement between scanners when assessed via Bland-Altman analysis. *D** and *K* were also unstable; with *D** being significantly greater at 3T compared to 1.5T and *K* being significantly higher on the 1.5T. To be defined as having a low variability CV and bias levels needed to be ≤ 15%. For Bland-Altman analysis *D** showed a bias of 24% and also had a high CV (78%). *K* was near to the threshold of low variability with a bias of 15% and a CV of 16%.

D_k and f showed good reproducibility with CV ≤ 15%, and α also showed low variability (CV: < 5%). ADC and D had very low variability (CV < 3%), with very little bias seen with Bland-Altman analysis (<1%).

For reproducibility of parameters based on different b values, *DDC* and *K* were the most unstable, showing no agreement between b value ranges and had high CV (*DDC*: 116%, *K*: 32%). The remainder of parameters showed small biases (≤ 5%) when assessed with Bland-Altman analysis.

The diffusion parameters ADC, D, D_k, and α all showed good reproducibility between different field strengths and b values. This is in line with the previously reported research[7,10–14,19–21]. Therefore, confidence can be placed in these parameters when measured in renal tissue, as variations in b values and field strength did not appear to affect these values.

D^* had high levels of variability when calculated between different field strengths. This is unsurprising when reviewing past literature[10–12,14,21], as D^* has been shown to be highly variable both when field strength was unchanged, and when it was altered. The instability of D^* has been attributed to the need for lower b values[10], as these have the greatest effects on this parameter[3].

Field strength may impact diffusion measurements for several reasons. Firstly, SNR increases linearly with field strength; in a 3T system the net magnetisation is twice that compared to a 1.5T system, which creates the potential for the signal to be twice as great[28,29]. However, due to various alterations of relaxation times and body heating, it is not exactly twice as great and is more likely 1.7 – 1.8 times greater[30,31]. Despite the increase in SNR, there are some disadvantages to imaging this higher field strength. 3T imaging compared to 1.5T increases magnetic susceptibility artefacts, which can give rise to loss of image quality[28,29,32]. These effects are doubled between 1.5T and 3T and cause image distortion and signal loss due to tissue boundaries, such as between soft-tissue and gases, which are common in the abdomen[28]. In the present study image quality was not assessed as the focus was on quantification of a variety of parameters, however, for certain clinical applications this may be an important consideration. Investigations into diffusion measurements in phantoms have indicated that ADC measurements are more uniform in the isocentre of the magnet and variability occurs with distance from this point[32]. This was seen more at 3T compared to 1.5T as these systems tended to have steeper spatial dependences.

Despite the aforementioned effects of field strength, in the present study the majority of parameters were still highly reproducible. No significant differences were seen in ADC, D, f, D_k, and α, all had CV ≤ 15%, and Bland-Altman analyses also indicted that

these parameters had good agreement when measured at 1.5T and 3T. It was only *D*, K* and *DDC* which had higher variability.

The variability seen in *DDC* (which was present when measured between field strengths and different b value ranges) was not evident in previous research[12–14]. However, the reason for the instability in the present study may be due to the fitting technique applied; there were no upper limits for *DDC*. Upper bounds can be applied to the stretched exponential model for example, 10 x 10^{-3} mm^2/s was set as the maximum value for *DDC* in it the first description of this model[4], however this value is still very high as free water diffusivity at body temperature is 3 x 10^{-3} mm^2/s, and so this value may be seen as a more appropriate upper boundary. In this present study it was decided that no upper bounds would be used, to prevent data hitting the maximum limit, this means that several voxels with extremely high fitted parameter values, which were not physically valid, may have skewed the distribution and increased the mean. If an upper bound was given then many voxels would have just reached this limit, thus not producing meaningful results. Table 6.7 showed that maximum values of *DDC* were greater in the 3T data compared to the 1.5T, this may have been due to the smaller voxel size in the 3T system. A larger voxel could contain a variety of diffusivities and thus, the mean value may decrease based on these lower values in this range, whereas in a smaller voxel there may be less variation and thus only these extreme values were captured. However, despite this difference the 1.5T values were still not physically valid. Thus, if this parameter is used in future without upper bounds then it may be better to use the median, as this produced more physically plausible values in both the 1.5T and 3T systems, and when different b values were used (Tables 6.3 and 6.4). The variability of *K* (which was moderate between scanners and high between b values) may also be explained by the lack of upper bounds, furthermore as previously mentioned *K* lacked reproducibility (CV > 30%) between two identical protocols measured 24 hours apart[14].

It should also be noted that this study used a voxel-by-voxel fitting method and then the mean value of all voxels was calculated within a ROI, as opposed to fitting the entire ROI. Therefore, variation from every voxel was included. This means that each voxel equally contributed to the mean providing a value reflective of the region under investigation. However, if a regional fitting method was used, then individual voxels with extreme values may have had less of an impact on the overall mean. This alternative approach may have resulted in lower, more physical valid, *DDC* results.

Past research has highlighted that *f* is less reproducible than *D*[11,12,14,21–23], and this was also demonstrated in this present study when compared between field strengths. However, in the present study *f* was reproducible (CV = 13.2%) when based on field

strength alterations and just above the threshold in the b value experiment (CV = 16.74%). Past research has cast doubt over the use of IVIM in the kidneys, suggesting that perfusion-related information cannot be obtained by this method[33]. However, more recent research has shown *f* to be clinically useful, with it being significantly related to renal function in those who had undergone partial nephrectomy[34]. Given the reasonable reproducibility in the current study and potential clinical utility, the evidence suggests that *f* should be considered for clinical use despite past concerns.

In this study the selection of b values had the greatest effect on *K* and *DDC* compared to the other parameters. As already stated, this may have been due to lack of upper bounds during the fitting process. Interestingly, both *D** and *f* had very little bias (< 5%) when assessed using Bland-Altman analysis based on b value alterations. This is interesting as optimum b value selection within the IVIM model is often discussed; the main finding being that a wide range of b values with many lower values are needed to get reproducible IVIM parameters (specifically *D** and *f*). This has been shown through simulations[35], in prostate cancer[36] and in renal tissue[35,37]. In the present study little variation was seen in *D** and *f* despite fewer low b values being part of the 7 b value acquisition. This may be due to the step-wise fitting technique of first fitting *D* at high b values ≥200 s/mm^2 and then fitting for *f* and *D*,* which is likely to have contributed to the stability of these parameters.

There were a few limitations with this study. Firstly, the sample size was small; to be able to generalise these findings and accurately determine which parameters are reproducible in renal tissue, a much larger cohort is needed. Secondly, voxel size differed between the 1.5T and 3T systems, which impacted the SNR and may have been a confounding variable in field strength comparisons. The acquisitions were matched to the clinical sequences used to assess paediatric abdominal tissue, and in these sequences different voxels sizes are used at different field strengths. The acquisitions were not altered as this study aimed to reflect clinical outcomes. Furthermore, these results can be used to assess how reliably clinically acquired DWI data could be combined from the different scanners at our institution. Therefore, although voxel size differences impacted on the results acquired at different field strength, it can be concluded that the majority of diffusion parameters under investigation were reproducible, when generated using optimised clinical protocols at 1.5T and 3T. This is promising for further investigations and clinical transference, as it highlights that many diffusion parameters do not greatly differ if they are generated with slight variations in acquisitions or field strength. Furthermore, although the effect of altering the range of b values was tested at 3T, only two sets of b values were investigated. Again, these sets were chosen to match those used

in clinical acquisitions at our institution, and in the 8 b value set additional measurements in the low-b-value range (<250 s/mm^2) were included, in order to increase sensitivity to the pseudo-random flow driven component of the DWI signal.

Additionally, the kidney ROIs were not separated into cortex and medulla. It has been previously shown that ADC values in the cortex are higher than in the medulla, regardless of b values used[38]. This may be related to the higher perfusion in the cortex, which can artificially elevate ADC in the mono-exponential model, which will also influence fitted values of IVIM parameters such as D^* and f in these regions. However, it is challenging to accurately separate these regions in the axial plane, particularly given the limited spatial resolution of the DWI acquisitions.

Finally, the left and right kidneys were not analysed separately. Previous work has shown D to be higher in the right kidney compared to the left in healthy volunteers and diabetic patients[39]. The reason for this may have been due to the gastrointestinal motility on the left side compared to the bowl motility on the right. Additionally, the right kidney is often slightly lower than the left due to the location of the liver. These slight physiological differences may have contributed to the differences found in D, however, the exact nature of why this parameter increased on the right side is unclear. The authors found no differences in D^* or f and no other models were tested, but it is possible that some of the variability seen in the present study may have been due to the left/right differences as kidneys were not analysed independently. However this variability is likely to have been a small factor and other research has failed to find significant differences in D, D^* and f between left and right kidneys[40]. Furthermore, the primary aim of this chapter was to investigate the reproducibility of fitted DWI parameters, rather than investigate left/right asymmetries in their absolute values.

6.6: Conclusions and Summary

Research into the use of DWI in the kidneys is growing and more advanced models of diffusion (non-Gaussian models) are starting to be used in the clinic. This study investigated the reproducibility of eight diffusion parameters from four models based on 5 pairs of healthy adult kidneys. The parameters were compared when measured at different field strengths (1.5T & 3T) and different b values.

The hypothesis was that ADC, D, DDC, α and D_k, would be reproducible in both experiments and that the greatest variability would be seen in D^* and f and K.

As predicted, ADC, D, α and D_k were shown to be reproducible in both experiments. In addition to these parameters, f was also found to be reproducible despite previous concerns over this parameter.

D^* and K were both found to have high levels of variability and were not reproducible in either experiment, this was also true for *DDC* which was not predicted.

These findings help place confidence in the clinical utility of certain parameters (ADC, D, f, D_k, and α) as they were reproducible despite acquisition alterations. This finding helps with generalising and interpreting results from other centres with different protocols. This study also suggested that more caution should be taken when interpreting D^*, K and *DDC* as these measurements lacked stability. This aids future research as certain findings may not be able to be replicated if different acquisitions are used.

Main Finding: Most parameters (ADC, D, D_k, and α) are reproducible across field strengths and when different b values are acquired. More caution needs to be taken when interpreting D^*, K and *DDC* as these parameters have higher levels of variability.

References: Chapter 6

1. McDonald, K., Sebire, N. J., Anderson, J. & Olsen, Ø. E. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy—feasibility study. *Pediatr. Radiol.* **41**, 99–106 (2011).
2. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
3. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
4. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
5. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
6. Grech-Sollars, M. *et al.* Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *Nmr Biomed.* **28**, 468–485 (2015).
7. Braithwaite, A. C., Dale, B. M., Boll, D. T. & Merkle, E. M. Short- and Midterm Reproducibility of Apparent Diffusion Coefficient Measurements at 3.0-T Diffusion-weighted Imaging of the Abdomen. *Radiology* **250**, 459–465 (2009).
8. Thoeny, H. C., De Keyzer, F., Oyen, R. H. & Peeters, R. R. Diffusion-weighted MR imaging of kidneys in healthy volunteers and patients with parenchymal diseases: initial experience. *Radiology* **235**, 911–917 (2005).
9. Cutajar, M., Clayden, J. D., Clark, C. A. & Gordon, I. Test-retest reliability and repeatability of renal diffusion tensor MRI in healthy subjects. *Eur. J. Radiol.* **80**, e263-268 (2011).
10. Jakab, A., Tuura, R., Kottke, R., Kellenberger, C. J. & Scheer, I. Intra-voxel incoherent motion MRI of the living human foetus: technique and test–retest repeatability. *Eur. Radiol. Exp.* **1**, (2017).
11. Sun, H., Xu, Y., Xu, Q., Shi, K. & Wang, W. Rectal cancer: Short-term reproducibility of intravoxel incoherent motion parameters in 3.0T magnetic resonance imaging. *Medicine (Baltimore)* **96**, e6866 (2017).

12. Reischauer, C., Patzwahl, R., Koh, D.-M., Froehlich, J. M. & Gutzeit, A. Non-Mono-Exponential Analysis of Diffusion-Weighted Imaging for Treatment Monitoring in Prostate Cancer Bone Metastases. *Sci. Rep.* **7**, (2017).
13. Winfield, J. M. *et al.* Modelling DW-MRI data from primary and metastatic ovarian tumours. *Eur. Radiol.* **25**, 2033–2040 (2015).
14. Jerome, N. P. *et al.* Repeatability of derived parameters from histograms following non-Gaussian diffusion modelling of diffusion-weighted imaging in a paediatric oncological cohort. *Eur. Radiol.* **27**, 345–353 (2017).
15. Pentang, G. *et al.* Diffusion kurtosis imaging of the human kidney: a feasibility study. *Magn. Reson. Imaging* **32**, 413–420 (2014).
16. Kwee, T. C., Takahara, T., Koh, D.-M., Nievelstein, R. A. J. & Luijten, P. R. Comparison and reproducibility of ADC measurements in breathhold, respiratory triggered, and free-breathing diffusion-weighted MR imaging of the liver. *J. Magn. Reson. Imaging* **28**, 1141–1148 (2008).
17. Choi, J. S. *et al.* Comparison of breathhold, navigator-triggered, and free-breathing diffusion-weighted MRI for focal hepatic lesions. *J. Magn. Reson. Imaging JMRI* **38**, 109–118 (2013).
18. Ye, X.-H., Gao, J.-Y., Yang, Z.-H. & Liu, Y. Apparent diffusion coefficient reproducibility of the pancreas measured at different MR scanners using diffusion-weighted imaging. *J. Magn. Reson. Imaging JMRI* **40**, 1375–1381 (2014).
19. Barral, M. *et al.* Diffusion-weighted MR imaging of the normal pancreas: Reproducibility and variations of apparent diffusion coefficient measurement at 1.5- and 3.0-Tesla. *Diagn. Interv. Imaging* **94**, 418–427 (2013).
20. Dale, B. M., Braithwaite, A. C., Boll, D. T. & Merkle, E. M. Field strength and diffusion encoding technique affect the apparent diffusion coefficient measurements in diffusion-weighted imaging of the abdomen. *Invest. Radiol.* **45**, 104–108 (2010).
21. Cui, Y., Dyvorne, H., Besa, C., Cooper, N. & Taouli, B. IVIM diffusion-weighted imaging of the liver at 3.0 T: Comparison with 1.5 T. *Eur. J. Radiol. Open* **2**, 123–128 (2015).
22. Barbieri, S., Donati, O. F., Froehlich, J. M. & Thoeny, H. C. Comparison of Intravoxel Incoherent Motion Parameters across MR Imagers and Field Strengths: Evaluation in Upper Abdominal Organs. *Radiology* **279**, 784–794 (2015).
23. Jakab, A., Tuura, R., Kottke, R., Kellenberger, C. J. & Scheer, I. Intra-voxel incoherent motion MRI of the living human foetus: technique and test–retest repeatability. *Eur. Radiol. Exp.* **1**, (2017).

24. Naganawa, S. *et al.* Diffusion-weighted Imaging of the Liver: Technical Challenges and Prospects for the Future. *Magn. Reson. Med. Sci.* **4**, 175–186 (2005).
25. Koh, D.-M. & Collins, D. J. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *Am. J. Roentgenol.* **188**, 1622–1635 (2007).
26. Huang, S. Y. *et al.* Body MR Imaging: Artifacts, k-Space, and Solutions. *Radiogr. Rev. Publ. Radiol. Soc. N. Am. Inc* **35**, 1439–1460 (2015).
27. Le Bihan, D. & Iima, M. Diffusion Magnetic Resonance Imaging: What Water Tells Us about Biological Tissues. *PLoS Biol.* **13**, (2015).
28. Chang, K. J., Kamel, I. R., Macura, K. J. & Bluemke, D. A. 3.0-T MR Imaging of the Abdomen: Comparison with 1.5 T. *RadioGraphics* **28**, 1983–1998 (2008).
29. Springer, F., Martirosian, P., Boss, A., Claussen, C. D. & Schick, F. Current Problems and Future Opportunities of Abdominal Magnetic Resonance Imaging at Higher Field Strengths. *Top. Magn. Reson. Imaging* **21**, 141–148 (2010).
30. Barth, M. M., Smith, M. P., Pedrosa, I., Lenkinski, R. E. & Rofsky, N. M. Body MR Imaging at 3.0 T: Understanding the Opportunities and Challenges. *RadioGraphics* **27**, 1445–1462 (2007).
31. Merkle, E. M. & Dale, B. M. Abdominal MRI at 3.0 T: The Basics Revisited. *Am. J. Roentgenol.* **186**, 1524–1532 (2006).
32. Malyarenko, D. *et al.* Multi-system repeatability and reproducibility of apparent diffusion coefficient measurement using an ice-water phantom. *J. Magn. Reson. Imaging JMRI* **37**, 1238–1246 (2013).
33. Müller, M. F., Prasad, P. V. & Edelman, R. R. Can the IVIM model be used for renal perfusion imaging? *Eur. J. Radiol.* **26**, 297–303 (1998).
34. Schneider, M. J. *et al.* Intravoxel Incoherent Motion Magnetic Resonance Imaging in Partially Nephrectomized Kidneys. *Invest. Radiol.* **51**, 323–330 (2016).
35. Zhang, J. L. *et al.* Optimization of b-value sampling for diffusion-weighted imaging of the kidney. *Magn. Reson. Med.* **67**, 89–97 (2012).
36. Pang, Y. *et al.* Intravoxel incoherent motion MR imaging for prostate cancer: An evaluation of perfusion fraction and diffusion coefficient derived from different b-value combinations. *Magn. Reson. Med.* **69**, 553–562 (2013).
37. Eisenberger, U. *et al.* Evaluation of renal allograft function early after transplantation with diffusion-weighted MR imaging. *Eur. Radiol.* **20**, 1374–1383 (2010).
38. Sulkowska, K. *et al.* Diffusion-weighted MRI of kidneys in healthy volunteers and living kidney donors. *Clin. Radiol.* **70**, 1122–1127 (2015).

39. Deng, Y. *et al.* Use of intravoxel incoherent motion diffusion-weighted imaging to detect early changes in diabetic kidneys. *Abdom. Radiol.* 1–6 (2018). doi:10.1007/s00261-018-1521-4

40. Mao, W. *et al.* Intravoxel incoherent motion diffusion-weighted imaging for the assessment of renal fibrosis of chronic kidney disease: A preliminary study. *Magn. Reson. Imaging* **47**, 118–124 (2018).

Chapter 7: Comparison of Models of Diffusion in Wilms' Tumour and Contralateral Renal Tissue

This chapter describes an original research study which explores the fitting of various diffusion models in Wilms' tumour and healthy renal tissue. This retrospective study compared the fits of four models of diffusion using the Akaike information criterion. It was found that in both Wilms' tumours and the contralateral normal renal tissue, non-Gaussian models provided superior fits compared to the mono-exponential model. Wilms' tumours were shown to be heterogenous with many different non-Gaussian models providing good fits across a single tumour.

7.1: Background: Comparing the Fits of Diffusion Models to Raw DWI data in the Body

It has previously been shown that ADC can be a very useful parameter in Wilms' tumour (see Chapter 4.2 for full details). Furthermore, although there is only limited research into the use of non-Gaussian models of diffusion in Wilms' (see Chapter 4.3), these models have the potential to provide additional information regarding tissue biology, as has been seen in renal tissue and other body pathologies (see Chapter 3 for full details). While previously the reproducibility of multiple diffusion parameters has been explored in healthy renal tissue (Chapter 6), it still needs to be established if these models provide a good representation of the raw diffusion data in Wilms' tumour.

There are many models which could be applied to multiple b value diffusion data, and so model selection is an important point of consideration. There are various methods for selecting the most appropriate model; such as R^2 selection, the Akaike information criterion (AIC)[1] and bayesian information criterion (BIC)[2]. A model's R^2 value refers to the coefficient of determination, and it describes the proportion of variance in the dependent variable that is predicted from the independent variable or variables. In terms of model selection, one could compare R^2 values from different models and then select the model with the highest R^2.

The aim of both AIC and BIC techniques is to determine which model best approximates the data[3]. Complex models with many fitted parameters can result in overfitting the data, and so both AIC and BIC minimise this problem by accounting for the

number of fitted parameters in the models. In both methods a value is produced per model, and it is the comparison of these values between different models which provides the relevant information. In both cases, the smallest AIC, or BIC value suggests the model is the most appropriate for the given data. The main difference between these two techniques is that BIC has a larger penalty for the number of fitted parameters compared to AIC.

Research into model comparisons has been explored in body imaging; in breast cancer[4], rectal cancer[5,6], prostate cancer bone metastases[7], ovarian cancer[8] and renal investigations[6,9–11]. The following sections summarise this literature; it can be seen that model preferences vary depending on the organ and pathology.

Mono-exponential and IVIM[12] models have been compared in breast cancer using AIC[4]. It was shown that out of 29 tumours IVIM was favoured by just over half (15) and mono-exponential was the preferred model for the remaining 14. As this difference is not great it suggests that both models produce adequate fits to the data. However, the more advanced invasive tumours tended to prefer the IVIM model, whereas the precursor tumours favoured the mono-exponential model. It was suggested that the presence of bi-exponential decay may indicate increased micro-vascularity which is seen in the more advanced tumours compared to the precursor tumours. However, this distinction in model preference was not seen in all tumours, and the numbers of each tumour group were very low, therefore without further analysis and an increased cohort this cannot be confirmed.

A slight preference for the mono-exponential model over IVIM was seen in rectal cancer[5]. Lesions were assessed based on which model the raw data in the majority of voxels was best described by, according to AIC. Out of 19 lesions, 12 showed a voxel-based majority for the mono-exponential model. However, there was high heterogeneity within the lesions, with many voxels also favouring the IVIM model. Thus, although the majority selected mono-exponential, a large proportion of the tumour was also well described by the IVIM model. In another study of rectal cancer[6], these models were again compared, along with the stretched exponential model[13]. Model comparisons were based on R^2 values and it was revealed that stretched exponential provided the best description of the data in both rectal cancer and healthy rectal tissue. IVIM had the smallest R^2 values (cancer tissue: 0.7, healthy tissue: 0.83) compared to mono-exponential (cancer tissue: 0.94, healthy tissue: 0.95) and stretched exponential (cancer tissue: 0.998, healthy tissue: 0.999), and provided a significantly poorer fit to the data.

In a study of patients with prostate cancer bone metastases, the three previously mentioned models were compared[7]. This study investigated the reproducibility of the parameters (see Chapter 6.1.2), but it also compared which model provided the best fit to the raw diffusion data. Model comparisons were based on BIC and it was shown that prior to and post- treatment, IVIM and stretched exponential were superior to mono-exponential. In 85% of the pre-treatment tumours non-Gaussian models were preferred, this was also found in 91-97% of post-treatment tumours (assessed at 3 time-points after treatment). Overall, the stretched exponential model was found to provide the best fit to the data in the majority of pre and post-treatment tumours.

An investigation into ovarian cancer also explored the mono-exponential, IVIM and stretched exponential models[8]. The IVIM model showed low repeatability for f and D^* and was excluded from AIC analysis, the two remaining models, however, were compared. It was shown that in pre-treatment patients the stretched exponential model was preferred in 64% of tumours, and in 65% of tumours post-treatment. The presence of an extra parameter provided by the stretched exponential model, its low variability, and its goodness of fit suggested that this model may be more useful in ovarian cancer compared to both mono-exponential and IVIM.

There is limited research into model comparisons using kurtosis[14] in body imaging. However, one feasibility study did demonstrate that the R^2 from the kurtosis model was higher than the R^2 from mono-exponential in healthy renal tissue[10]. However, this was only a marginal improvement (mono-exponential: $R^2 = 0.96$, kurtosis: $R^2 = 0.99$). Therefore, much further research is needed to determine the goodness of fit of kurtosis modelling compared to other diffusion models.

From the described research it seems that the stretched exponential model provides a good fit to diffusion data in a variety of organs. The IVIM model, however, appears to be more variable and does not always describe the data better than a mono-exponential fit. When it comes to renal imaging, however, IVIM does seem superior to mono-exponential[9,11]. Diffusion values within the renal cortex and medulla vary according to different studies and thus an investigation into mono-exponential fitting sought to determine whether the model was, in part, responsible[11]. By acquiring data in three healthy volunteers and resampling the signal decay curves based on different b values, it was shown that variation in diffusivity coefficients was largely due to the choice of b values, and that the signal decay was bi-exponential and not mono-exponential. The authors suggested that future work should use the bi-exponential model as this provided a better description of the diffusion signal in healthy renal tissue. Furthermore, they

highlighted that IVIM produced more reproducible values of *D* compared to ADC[11], and the additional parameters *D** and *f* may provide further insight into kidney status and function, as these parameters potentially reflect vascular perfusion and tubular flow[15].

Another study also showed that in healthy renal tissue IVIM was superior than a mono-exponential model[9]. It was shown that <8% of pixels from whole kidney volumes of five volunteers were best described by the mono-exponential model. Therefore, in healthy renal tissue, it appears that IVIM may be the best descriptor of diffusion data.

7.2: Aims and Hypotheses

Mono-exponential, IVIM, stretched exponential and kurtosis have the potential to be useful in Wilms' tumour. All of these models provide diffusion parameters which have shown promise in a variety of organs including the kidneys. The following investigation sought to determine which models best described the raw diffusion signal in Wilms' tumours and the contralateral unaffected renal tissue. It was expected that IVIM may be superior to mono-exponential in the unaffected renal tissue based on previous research[9,11]. Within Wilms' tumour, however, no predications were made as there is little literature on this topic in this specific pathology and no clear model preferences have been shown in other body cancers[4,5,7,8,16]. However, it was expected that as Wilms' tumour subtypes have different cellular environments, different models may favour different subtypes.

7.3: Methods

7.3.1: Study Population

Institutional ethical approval was granted for all studies within this thesis. A 10-year retrospective review (April 2007 – March 2017) of the radiology imaging system at Great Ormond Street Hospital, London, UK, was performed for all abdominal MRI data in children with a proven histological diagnosis of Wilms' tumour. Inclusion criteria were those with multiple b value DWI data (with a maximum b value of 1,000 s/mm^2), and tumour size covering at least 2 axial slices on DWI. Data was collected from Wilms' tumour patients both pre- and post-chemotherapy. For normal kidney data, the contralateral unaffected kidney was used.

7.3.2: MRI

All imaging was performed on a 1.5T Siemens Magnetom Avanto scanner equipped with 40 mT/m gradients. Depending on patient size, one or two body matrix coils were used to obtain full coverage (6 element design, Siemens). Patients were either awake or anaesthetised depending on their age.

Multiple b value DWI was obtained for all patients and was acquired during free breathing. The DWI protocol was as follows: 7 or 8 b values in 3 orthogonal directions (0, 50, 100, 250, 500, 750, 1000 s/mm^2 or 0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2) slice thickness: 6mm, TR/TE: 2800ms/89ms, field of view: 350mm^2, voxel size: 1.4 x 1.4 x 6mm, number of slices: 19, matrix size: 128 x 96 x 19. Each b value had 9 averages and trace images were used for analysis.

7.3.3: Processing and ROIs

Diffusion data was processed four times using the trace images from each Wilms' tumour patient using four models of diffusion: mono-exponential, IVIM, stretched exponential, and kurtosis. The details of these fitting methods are described in Chapter 6.3.3.

ROIs were generated using Mango Software (Research Imaging Institute, UTHSCSA). ROIs were drawn on the b0 images around the entire tumour volume, these were edited and verified by a radiologist specialising in paediatric radiology (2 years' experience in paediatric radiology). Normal kidney tissue was also defined on the b0 images and areas of high flow such as the areas which surround the renal pelvis were excluded; an example can be seen in Figure 7.1. All analysis regarding model comparisons was confined to these ROIs.

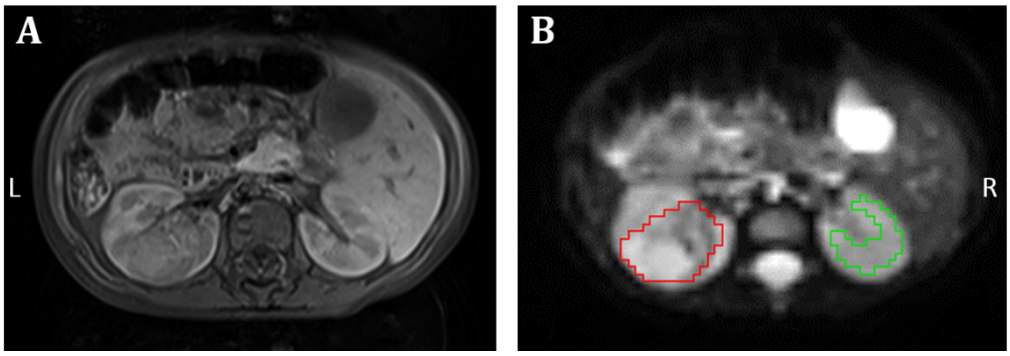

Figure 7.1. An example of a representative Wilms' tumour. Displayed is a central axial slice of a T_1w image **(A)** and b0 image **(B)**. Shown is the abdomen at the level of the kidney of a Wilms' tumour patient post-chemotherapy (age at time of scan: 1.22 years). ROIs are shown surrounding the tumour (red) and normal renal tissue (green). In this diagram the right side of the image (white R) is the patients' right side.

7.3.4: Model Comparison Analysis

AIC was used to compare the four models. For every voxel within the tumour ROIs and normal kidney ROIs, AIC was calculated for each model. The model with the lowest AIC value within that voxel was classified as the 'winner'. The voxel was then labelled based on which model 'won'; 1 = mono-exponential, 2 = IVIM, 3 = stretched exponential, 4 = kurtosis. This process was repeated across the entire ROI volumes. The modal value was then calculated per tumour and normal kidney tissue. This value therefore represented the model which provided the best overall fit for the tissue. A diagram for these methods can be seen in Figure 7.2. Additionally, models were compared between Wilms' tumour subtypes to determine whether a certain subtype favoured a particular model.

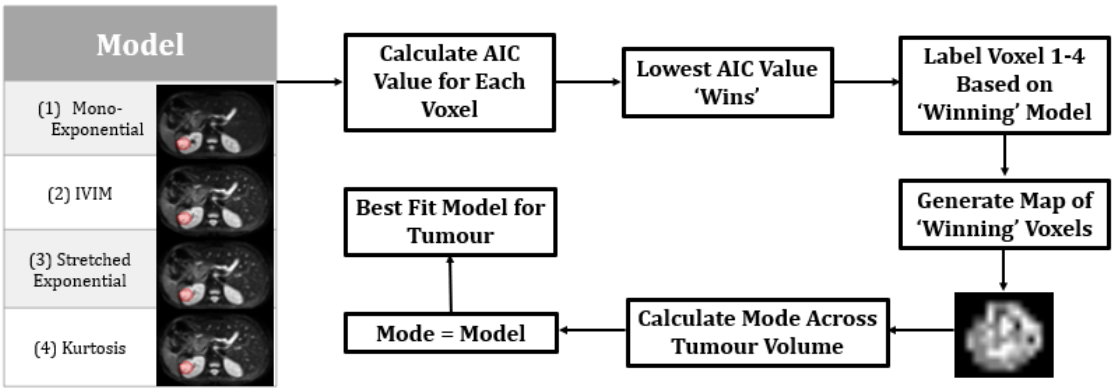

Figure 7.2. Diagram detailing methodology for comparing models. AIC = Akaike information criterion

7.4: Results

7.4.1: Study Population

A total of 50 pre-chemotherapy and 62 post-chemotherapy Wilms' tumours were included for diffusion model comparison analysis. A flow chart detailing reasons and numbers of excluded cases can be seen in Figure 7.3. The mean age of patients at their pre-chemotherapy scan was 2.5 years (*SD*: 2.2), and the mean age at their post-chemotherapy scan was 3.1 years (*SD*: 2.7).

The diffusion data was acquired from either 7 or 8 b values (0, 50, 100, 250, 500, 750, 1000 s/mm^2 or 0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2). Fifty-one tumours had the 7 b value protocol (23 pre-chemotherapy and 28 post-chemotherapy), and 61 tumours had the 8 b value protocol (27 pre-chemotherapy and 34 post-chemotherapy).

Of the 62 post-chemotherapy tumours, 57 had histologically confirmed subtypes: 7 blastemal, 9 epithelial, 13 stromal, 8 regressive, 19 mixed and 1 completely necrotic. Subtypes were defined according to SIOP-2001 protocol, for details on this process please see Chapter 1.5.

The contralateral unaffected kidney was used as the normal kidney data, due to bilateral cases this gave a total of 77 normal kidney datasets; 39 from patients who had received chemotherapy and 38 from patients who had not. Of the 77 normal kidney datasets 33 had the 7 b value protocol (16 pre- chemotherapy and 17 post-chemotherapy), and 44 had the 8 b value protocol (22 pre- and 22 post-chemotherapy).

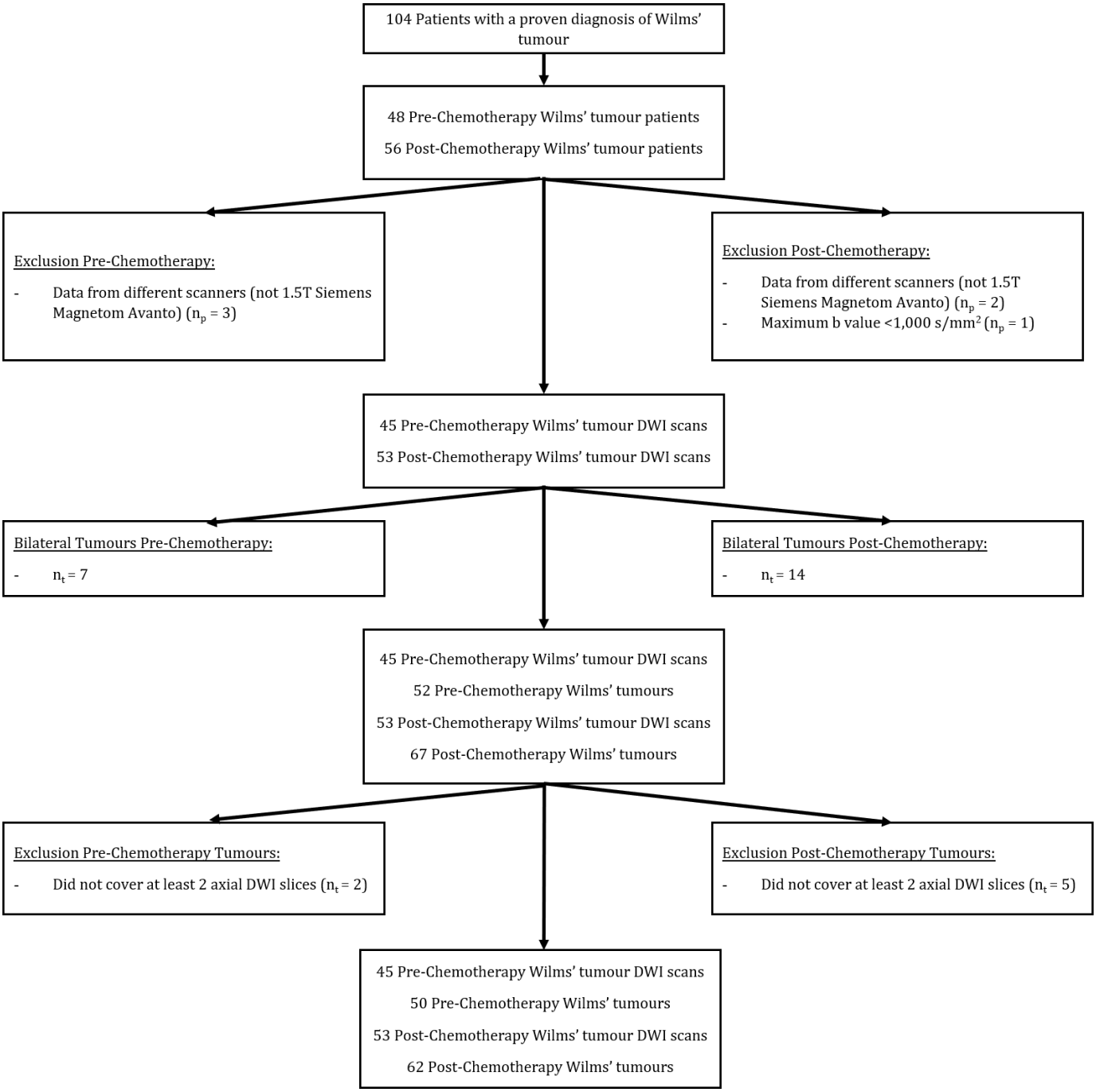

Figure 7.3. Flowchart of study population showing exclusion criteria. DWI = diffusion weighted imaging. n_p = number of patients, n_t = number of tumours.

7.4.2: Wilms' Tumour AIC Results

Combining both the pre- and post-chemotherapy datasets it was shown that the stretched exponential model provided the best fit to the majority of the raw diffusion data. Of the 112 Wilms' tumours analysed, the stretched exponential model provided the best fit to 61/112 of the tumours (54.5%). The kurtosis model provided the best fit to 29/112 (25.9%), IVIM provided the best model for 21/112 (18.8%), and mono-exponential provided the best fit to 1/112 of the Wilms' tumours (0.9%).

When the data was separated by pre-and post-chemotherapy, a similar pattern appeared. For pre-chemotherapy data stretched exponential provided the best fit for the majority of the data: 31/50, (62%). Kurtosis provided the best fit for 14/50 (28%), IVIM for 5/50 (10%) and mono-exponential did not provide the best fit for any of the pre-chemotherapy Wilms' tumour data.

For the post-chemotherapy Wilms' tumour data, stretched exponential was the most common best fit model: 30/62 (48.4%). IVIM provided the best model for 16/62 (25.8%) and Kurtosis provided the best model for 15/62 (24.2%). Mono-exponential provided the best fit to 1/62 of the post-chemotherapy Wilms' tumours (1.6%).

As it could be suggested that the number of b values might have affected the fits to the diffusion data, tumours were separated into 7 and 8 b value groups. Overall, stretched exponential was the preferred model for the 7 b value group: 22/51 (43.2%). Kurtosis provided the best fit for 16/51 (31.4%), and IVIM for 13/51 (25.5%). Mono-exponential did not provide the best fit for any of the 7 b value tumours.

For the 8 b value Wilms' tumour data, a similar pattern emerged: stretched exponential provided the best fit for the majority of the tumours: 39/61 (63.9%). Kurtosis provided the best for 13/61 (21.3%), IVIM for 8/61 (13.1%), and mono-exponential was the best fit for one of the 8 b value tumours (1.6%).

Table 7.1 shows the comparisons based on b values and whether or not chemotherapy was administered, in each instance stretched exponential was the preferred model. Figure 7.4 provides an example of the model fits to the raw diffusion data in a post-chemotherapy 7 b value Wilms' tumour, and similar patterns were seen for the untreated and 8 b value tumours.

It was also tested to see if the best fit model was related to Wilms' tumour subtype. For subtypes which have been confirmed post-surgery, the best fit models were compared using the post-chemotherapy data (nearest time point to histology). However, it was

shown that models did not particularly favour certain subtypes; with the stretched exponential model performing well across most subtypes. Full details can be seen in table 7.2.

Table 7.1. Model comparisons based on Akaike information criterion for Wilms' tumours, separated by number of b values used in the diffusion acquisition, and whether or not chemotherapy had been administered. The percentages refer to the percent of tumours which favoured each model

	Pre-chemotherapy		Post-chemotherapy	
	7 b values n = 23	8 b values n = 27	7 b values n = 28	8 b values n = 34
Mono-exponential	0%	0%	0%	2.9%
IVIM	21.7%	0%	28.6%	23.5%
Stretched exponential	52.2%	70.4%	35.7%	58.8%
Kurtosis	26.1%	29.6%	35.7%	14.7%

Table 7.2. Model comparisons based on Akaike information criterion per tumour subtype. The percentages refer to the percent of tumours which favoured each model

	Blastemal (n = 7)	Epithelial (n = 9)	Stromal (n = 13)	Regressive (n = 8)	Mixed (n = 19)	Necrotic (n =1)
Mono-exponential	0%	0%	0%	12.5%	0%	0%
IVIM	28.6%	11.1%	15.4%	37.5%	26.3%	0%
Stretched exponential	71.4%	77.8%	46.2%	25%	36.8%	100%
Kurtosis	0%	11.1%	38.5%	25%	36.8%	0%

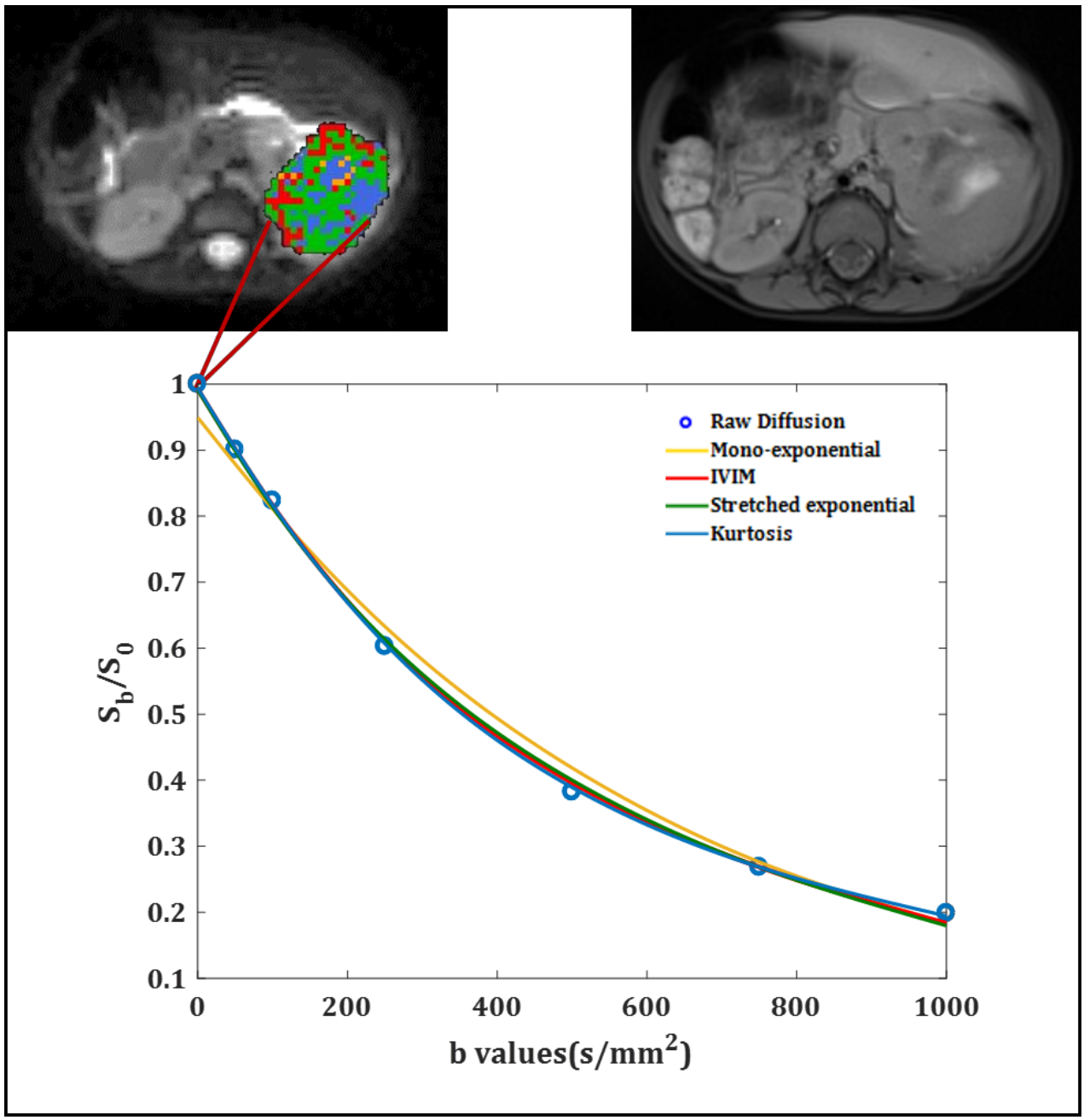

Figure 7.4. A central axial slice of a b0 image (left) and the clinical T_1w image (right) is shown of the abdomen at the level of the kidney. This is a representative post-chemotherapy patient (Age: 2.44 years). The Wilms' tumour ROI is highlighted and is colour-coded according to which model provided the best fit to the raw diffusion data, based on the Akaike information criterion; yellow = mono-exponential, red = IVIM, green = stretched exponential, blue = kurtosis. The graph shows the fits of four models to the raw diffusion data across the ROI volume. Normalised signal is shown on the y axis (S_b = signal at a given b values, S_0 = signal where b = 0) and b values are shown on the x axis.

7.4.3: Normal Kidney AIC Results

The normal kidney data (contralateral unaffected kidney) showed that IVIM provided the best model for the majority of the data: 47/77 (61%). Stretched exponential was the only other model which provided the best fit: 30/77(39%).

Even though normal kidney tissue was being investigated, the data was separated into pre- and post-chemotherapy groups in case the treatment affected the normal kidney. There was a slight discrepancy between the pre- and post-chemotherapy scans. For pre-chemotherapy data half of the contralateral kidneys favoured IVIM as the best model and the other half favoured stretched exponential (19/38, 50%). Whereas in the post-chemotherapy data IVIM provided the best model for the majority of the contralateral kidneys (28/39, 71.8%), and stretched exponential provided the best fit for the remaining data (11/39, 28.2%).

The normal kidney data was also separated into b values, and it was shown that stretched exponential was favoured by the 7 b value data in the majority of contralateral kidneys: 19/33 (57.6%), with IVIM being favoured by the remaining contralateral kidneys: 14/33 (42.4%). For the 8 b value data IVIM provided the best fit to the majority of the contralateral kidneys: 33/44 (75%), and stretched exponential was the best model for the remaining contralateral kidneys: 11/44 (25%).

Table 7.3 shows the comparisons based on b values and whether or not chemotherapy was administered. It can be seen that regardless of whether the patients had undergone treatment, the 8 b value range favoured the IVIM model. However, the 7 b value range showed a preference for stretched exponential prior to treatment and a slight preference for IVIM post-treatment. Figure 7.5 and 7.6 show examples of the model fits to the raw diffusion data for both a 7 b value and 8 b value pre-chemotherapy normal kidney.

Table 7.3. Model comparisons based on Akaike information criterion for contralateral normal kidney tissue, separated by number of b values used in the diffusion acquisition, and whether or not chemotherapy had been administered. The percentages refer to the percent of tumours which favoured each model

	Pre-chemotherapy		Post-chemotherapy	
	7 b values n = 16	8 b values n = 22	7 b values n = 17	8 b values n = 22
Mono-exponential	0%	0%	0%	0%
IVIM	31.3%	63.6%	52.9%	86.4%
Stretched exponential	68.8%	36.4%	47.6%	13.6%
Kurtosis	0%	0%	0%	0%

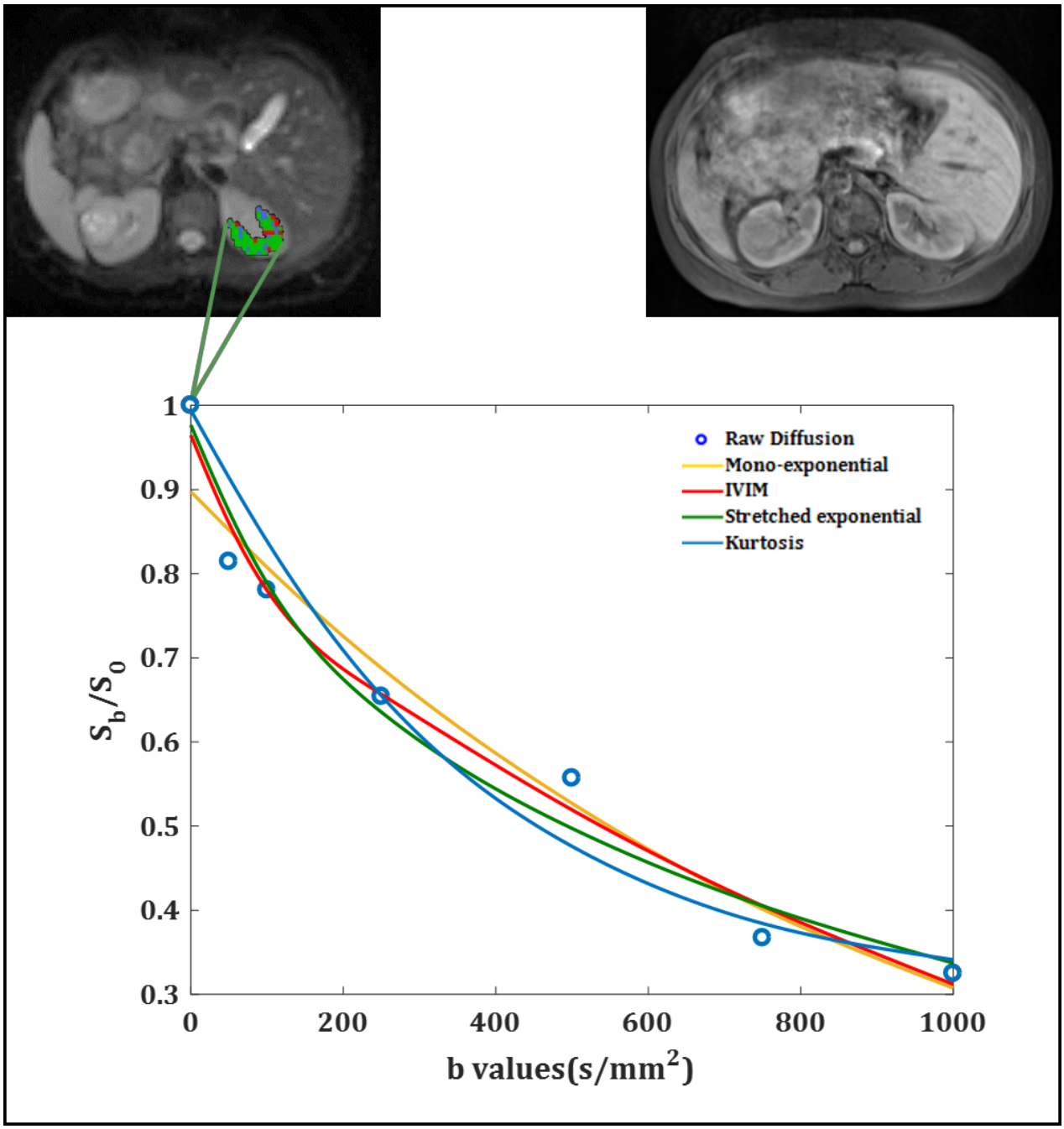

Figure 7.5. A central axial slice of a b0 image (left) and the clinical T_1w image (right) is shown of the abdomen at the level of the kidney, the normal kidney ROI is shown in green. The ROI is highlighted and is colour-coded according to which model provided the best fit to the raw diffusion data, based on the Akaike information criterion; yellow = mono-exponential, red = IVIM, green = stretched exponential, blue = kurtosis. This is a representative pre-chemotherapy patient (Age: 2.8 years) with a 7 b value acquisition. The graph shows the fits of four models to the raw diffusion data across the ROI volume. Normalised signal is shown on the y axis (S_b = signal at a given b values, S_0 = signal where b = 0) and b values are shown on the x axis.

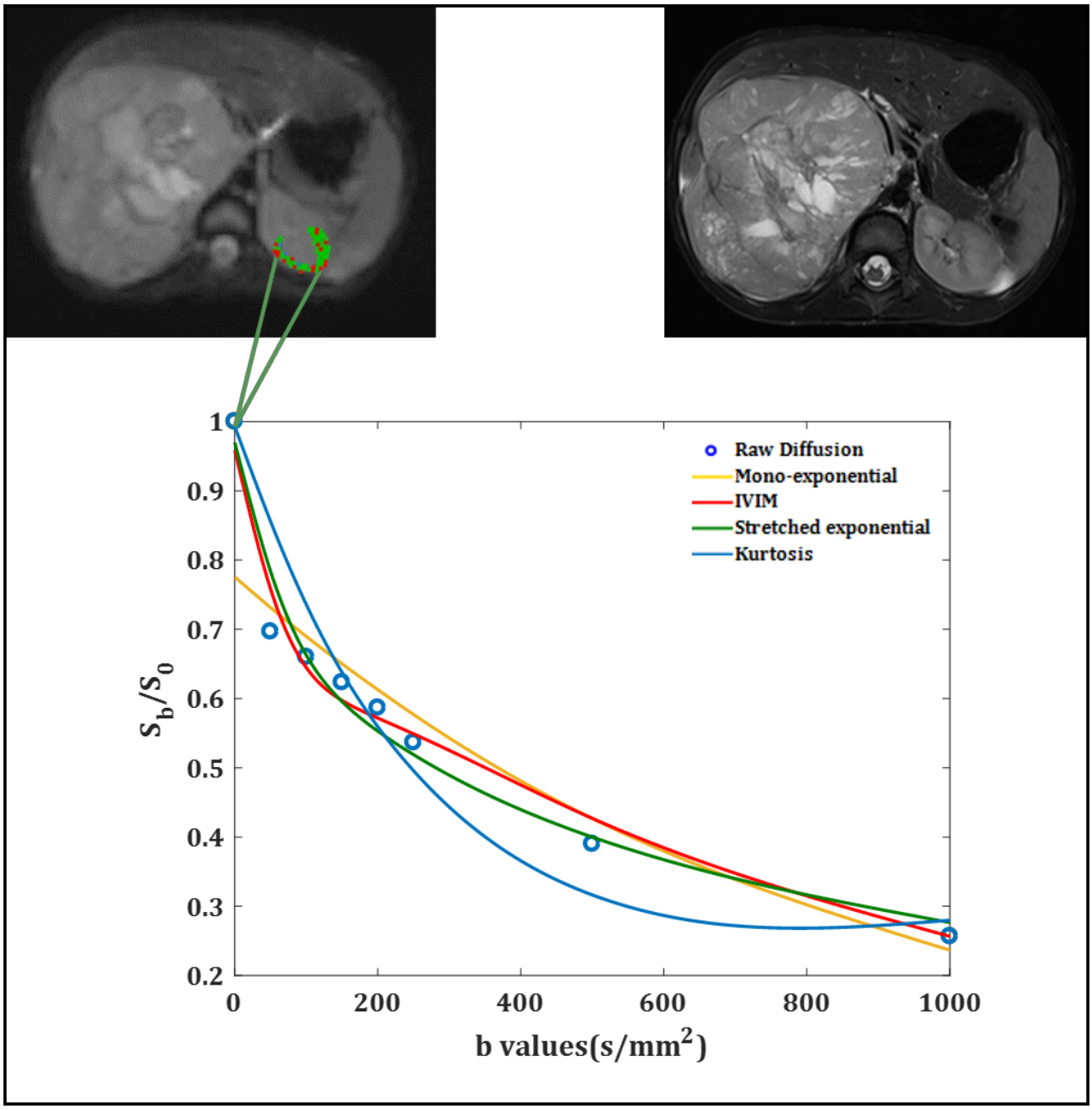

Figure 7.6. A central axial slice of a b0 image (left side) and the clinical T_1w image (right side) is shown of the abdomen at the level of the kidney, the normal kidney ROI is shown in green. The ROI is highlighted and is colour-coded according to which model provided the best fit to the raw diffusion data, based on the Akaike information criterion; yellow = mono-exponential, red = IVIM, green = stretched exponential, blue = kurtosis. This is a representative pre-chemotherapy patient (Age: 4.47 years) with an 8 b value acquisition. The graph shows the fits of four models to the raw diffusion data across the ROI volume. Normalised signal is shown on the y axis (S_b = signal at a given b values, S_0 = signal where b = 0) and b values are shown on the x axis.

7.5: Discussion

This study compared four models of diffusion based on how well they fit to the raw DWI data, according to AIC. These comparisons were made in Wilms' tumours, both pre- and post-chemotherapy, and on the contralateral unaffected kidney, as a measure of normal renal tissue. The diffusion data came from both 7 and 8 b value ranges. For the Wilms' tumour datasets, it was shown that the stretched exponential model provided the best fit to the majority of the data. This result was maintained in both b value groups, and in both treatment groups. There was no particular model preference when the tumours were investigated according to subtype. Mono-exponential was shown to be the least appropriate model according to AIC, with only one post-chemotherapy tumour favouring this model. Overall, the normal kidney tissue favoured the IVIM model, however when this data was separated into b value ranges, it was shown that stretched exponential was slightly favoured by the 7 b value range, and IVIM by the 8 b value range. No other models were found to be appropriate for the normal kidney data.

The main finding from this investigation was that non-Gaussian models provided better descriptions of the diffusion data compared to the mono-exponential model in Wilms' tumour and normal renal tissue. This finding was observed in all of the normal renal tissue data, and all of the pre-chemotherapy Wilms' tumour data. Only a single post-chemotherapy Wilms' tumour favoured the mono-exponential model, thus highlighting that it is very uncommon for diffusion signal to exhibit mono-exponential behaviour in normal renal tissue and Wilms' tumours. The deviation from a mono-exponential decay has been previously highlighted and explored, it has been shown that there was a rapid decline in signal at lower b values followed by a more gradual decline at higher b values in the liver[17]. This initial decline was suggested to be due to vascular perfusion, and thus lower b values are thought to be sensitive to signal attenuation from perfusion[12], giving rise to the IVIM model. This has been shown to be the case in healthy renal tissue[9,11], where signal was shown to be bi-exponential as opposed to mono-exponential; as the kidney is a well-perfused organ.

No particular subtype appeared to favour a certain model, however, the numbers in each group were small. Additionally, it is important to remember that subtypes are defined based on only analysing a subsection of the tumour. Wilms' tumours are very heterogeneous and across a single tumour or even slice, there will be areas of distinct cellular environments. If these tumours had been analysed on a slice-by-slice basis,

indicating regions of different cellular subtypes then they could have been matched to the MRI scans. In this instance it may have then been possible to see that the voxels which favoured certain models belonged to certain subtypes. However, due to the lack of advanced histology this was not possible and thus it appeared that subtypes had no model preference.

While the mono-exponential model did not provide the best fit to the DWI data, it does not mean that it should not be used clinically. In Wilms' tumour ADC has been useful in distinguishing between benign from malignant tumours[18,19], investigating chemotherapy-induced changes[20] and analysing subtypes[21,22]. Furthermore, ADC values have been shown to be reproducible between a 1.5T and 3T scanner (coefficient of variation [CV]: 1.03%) and when calculated based on different b value ranges (CV: 11.05%), when measured in healthy adult renal tissue (see Chapter 6 for full details). Additionally, ADC does not require multiple b values, which is beneficial as many centres may not acquire DWI with multiple b values as standard. Therefore, despite the present study showing a deviation from a mono-exponential signal decay, it is important to be aware that while the model may not be the best descriptor of the DWI data, it is nonetheless reproducible and clinically useful.

The stretched exponential model provided the best fit to the DWI Wilms' tumour data. The previously mentioned studies in rectal cancer and healthy rectal tissue[16], prostate cancer bone metastasis[7] and ovarian cancer[8], all showed this model to provide the best fit to DWI data when compared to IVIM and mono-exponential. The stretched exponential model provides two parameters, α and *DDC*; while the exact physiological basis of α is unknown, it is thought to represent tissue heterogeneity, with a lower value suggesting a more heterogenous environment[13]. As tumour tissue is very heterogenous, it is unsurprising that the stretched exponential model describes this data well. Previous work has shown both α and *DDC* to be reproducible[7,8], however *DDC* was found to have high variability in healthy kidneys (CV: 108 – 116%), (Chapter 6). This finding was likely to be due to fitting methods, and a lack of boundary conditions (for full details see Chapter 6.5). Therefore, despite the goodness of fit according to AIC, the fitted parameter *DDC* should be used with caution as values may just converge on the boundary limit or be highly variable. The α parameter however, has the potential to be a useful parameter in Wilms' and normal renal tissue, as it has lower variability (CV: 5% in healthy renal tissue, Chapter 6) and is generated from a model which describes the data well.

The IVIM model was favoured over the four other models by normal renal tissue; this is supported by previous work where IVIM was preferred over mono-exponential in

healthy renal tissue[9,11]. When the data was separated into b value ranges there was a slight preference for the 7 b value range to favour the stretched exponential model, and the 8 b value range to favour IVIM. The reason for this may be due to inclusion of lower b value weightings in the 8 b value range (0, 50, 100, 150, 200, 250, 500, 1000 s/mm^2), compared to the 7 b value range, which focused on higher diffusion weightings (0, 50, 100, 250, 500, 750, 1000 s/mm^2). Lower b values are needed to capture the fast pseudo random diffusion/perfusion effects[23], and thus the 8 b value range is more suited to an IVIM model than the 7 b value range.

The normal kidney data was separated in the pre- and post-treatment groups as although the treatment is not intended for this tissue, chemotherapy affects the entire body and thus would likely influence the contralateral kidney. Only a slight change in model selection was seen between these groups, with the pre-treatment group showing equal preference for either IVIM or stretched exponential, whereas in the post-treatment group 72% of the kidneys favoured IVIM. This may be due to the effects of chemotherapy changing the tissue. However, this is only a minor change in model selection and when the pre-treatment group is separated into b value ranges, majority of kidneys in the 8 b value group favoured IVIM as well. Therefore, it is unlikely that the treatment had a great effect on model selection in the normal renal tissue.

The method for selecting the model which provided the best fit is a potential weakness of this study. AIC takes into account the complexity of the model and goodness of fit, and therefore seemed an appropriate choice for model comparison and selection. However, this process was completed on a voxel-by-voxel basis and the model which most frequently 'won' across the entire tissue volume was chosen. This does not take into account the fact that the best fit model may have only been the best in a few more voxels than the next best fit model. It is important to consider if one model is clearly the best for the entire tissue or if there is only a small difference between the models. This was previously highlighted by Manikis et al.[5] in rectal cancer, where although overall mono-exponential was preferred to IVIM, there was high heterogeneity across the tissue. An example of this can be seen in Figure 7.7. It shows a pre-chemotherapy Wilms' tumour which favoured the IVIM model, however, it can be seen that there was very little difference between the number of voxels selecting stretched exponential, and the number of voxels selecting IVIM, (difference of 103 voxels). With this in mind, one should be cautious before claiming a particular model best fits the data, as it may be that many models are near equal in fitting quality. This once again highlights the highly heterogenous nature of Wilms' tumours. Thus, it is important to always be mindful when making

conclusions about Wilms's tumour tissue, whether it is in terms of model preference, necrotic regions or subtype.

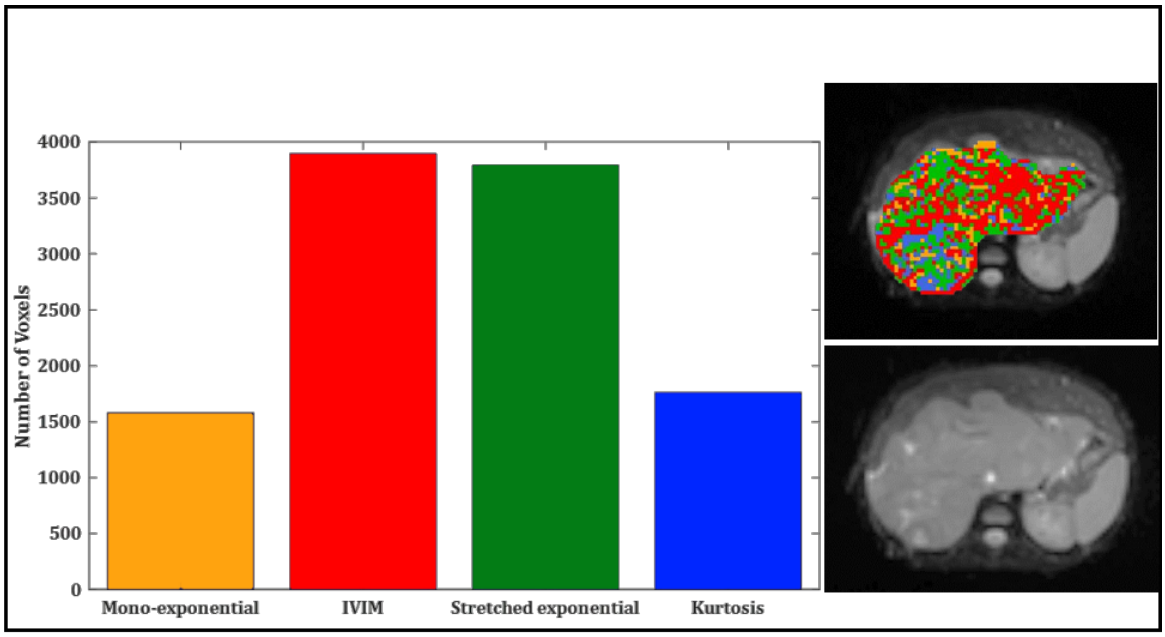

Figure 7.7. An example of a central axial slice (b0) of the abdomen. Each voxel within the Wilms' tumour (pre-chemotherapy tumour, age at scan: 6.45 years) is colour-coded according to which model provided the best fit to the raw diffusion data, based on the Akaike Information Criterion; yellow = mono-exponential, red = IVIM, green = stretched exponential, blue = kurtosis. A graph is shown detailing the number of voxels assigned to each model across the entire tumour volume.

7.6: Conclusions and Summary

Non-Gaussian models of diffusion have the potential to provide additional clinical information as opposed to standard ADC. This study investigated how well non-Gaussian and mono-exponential models describe DWI data, according to AIC, in Wilms' tumours and normal renal tissue. It was shown that the mono-exponential model does not fit the DWI data as well as IVIM, stretched exponential or kurtosis in Wilms' tumour tissue and normal renal tissue. Stretched exponential was the best model for describing the Wilms' data, but specific subtypes did not show a preference for different models. It is important to acknowledge that Wilms' tumours are very heterogenous and that within a single tumour many non-Gaussian models provided good descriptions of the data. As expected in the contralateral normal renal tissue IVIM was the best descriptor of the DWI data.

ADC is frequently used in clinical research and therefore the assumption is that the signal decay is mono-exponential. However, these results suggest that in Wilms' tumour and normal renal tissue, the DWI signal does not exhibit a mono-exponential decay. Therefore, utilising other models may provide more accurate representations of the underlying tissue environment. Thus, potentially improving our interpretation of research and our understanding of what the MRI signal is reflecting. However, one does need to keep in mind that these fits may be dependent on the acquisition, for example number of b values, and it is also important to consider if the fitted parameters are reproducible and clinically useful.

Main Finding: Non-Gaussian models provide superior fits to diffusion data in Wilms' tumours and contralateral normal kidney tissue compared to a mono-exponential model. Out of 112 tumours only 1 favoured the mono-exponential model, and out of 77 contralateral kidneys none favoured the mono-exponential model.

References: Chapter 7

1. Akaike, H. A new look at the statistical model identification. *IEEE Trans. Autom. Control* **19**, 716–723 (1974).
2. Schwarz, G. Estimating the Dimension of a Model. *Ann. Stat.* **6**, 461–464 (1978).
3. Frank J. Fabozzi, Sergio M. Focardi, Svetlozar T. Rachev & Bala G. Arshanapalli. Model Selection Criterion: AIC and BIC. in *The Basics of Financial Econometrics: Tools, Concepts, and Asset Management Applications.* 399–403 (John Wiley & Sons, Inc., 2014).
4. Yuan, J. *et al.* Statistical assessment of bi-exponential diffusion weighted imaging signal characteristics induced by intravoxel incoherent motion in malignant breast tumors. *Quant. Imaging Med. Surg.* **6**, 418–429 (2016).
5. Manikis, G. C. *et al.* Diffusion weighted imaging in patients with rectal cancer: Comparison between Gaussian and non-Gaussian models. *PLoS ONE* **12**, (2017).
6. Zhang, G. *et al.* Comparison of non-Gaussian and Gaussian diffusion models of diffusion weighted imaging of rectal cancer at 3.0 T MRI. *Sci. Rep.* **6**, 38782 (2016).
7. Reischauer, C., Patzwahl, R., Koh, D.-M., Froehlich, J. M. & Gutzeit, A. Non-Mono-Exponential Analysis of Diffusion-Weighted Imaging for Treatment Monitoring in Prostate Cancer Bone Metastases. *Sci. Rep.* **7**, (2017).
8. Winfield, J. M. *et al.* Modelling DW-MRI data from primary and metastatic ovarian tumours. *Eur. Radiol.* **25**, 2033–2040 (2015).
9. Wittsack, H.-J. *et al.* Statistical evaluation of diffusion-weighted imaging of the human kidney. *Magn. Reson. Med.* **64**, 616-622 (2010). doi:10.1002/mrm.22436
10. Pentang, G. *et al.* Diffusion kurtosis imaging of the human kidney: a feasibility study. *Magn. Reson. Imaging* **32**, 413–420 (2014).
11. Zhang, J. L. *et al.* Variability of renal apparent diffusion coefficients: limitations of the monoexponential model for diffusion quantification. *Radiology* **254**, 783–792 (2010).
12. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
13. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
14. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).

15. Thoeny, H. C. *et al.* Functional evaluation of transplanted kidneys with diffusion-weighted and BOLD MR imaging: initial experience. *Radiology* **241**, 812–821 (2006).
16. Zhang, G. *et al.* Comparison of non-Gaussian and Gaussian diffusion models of diffusion weighted imaging of rectal cancer at 3.0 T MRI. *Sci. Rep.* **6**, (2016).
17. Koh, D.-M. & Collins, D. J. Diffusion-Weighted MRI in the Body: Applications and Challenges in Oncology. *Am. J. Roentgenol.* **188**, 1622–1635 (2007).
18. Humphries, P. D., Sebire, N. J., Siegel, M. J. & Olsen, Ø. E. Tumors in Pediatric Patients at Diffusion-weighted MR Imaging: Apparent Diffusion Coefficient and Tumor Cellularity. *Radiology* **245**, 848–854 (2007).
19. Gawande, R. S., Gonzalez, G., Messing, S., Khurana, A. & Daldrup-Link, H. E. Role of diffusion-weighted imaging in differentiating benign and malignant pediatric abdominal tumors. *Pediatr. Radiol.* **43**, 836–845 (2013).
20. McDonald, K., Sebire, N. J., Anderson, J. & Olsen, O. E. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy-feasibility study. *Pediatr. Radiol.* **41**, 99–106 (2011).
21. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
22. Littooij, A. S. *et al.* Apparent diffusion coefficient as it relates to histopathology findings in post-chemotherapy nephroblastoma: a feasibility study. *Pediatr. Radiol.* **47**, 1608–1614 (2017).
23. Koh, D.-M., Collins, D. J. & Orton, M. R. Intravoxel Incoherent Motion in Body Diffusion-Weighted MRI: Reality and Challenges. *Am. J. Roentgenol.* **196**, 1351–1361 (2011).

Chapter 8: An Alternative Approach to Contrast-Enhanced Imaging: Diffusion Weighted Imaging and T_1-weighted Imaging Identifies and Quantifies Necrosis in Wilms' Tumour

This chapter describes a retrospective study which aimed to identify necrotic tissue in Wilms' tumours, without using contrast-enhanced imaging, but rather combining DWI and T_1w. It found that combining ADC and T_1w, maps could be created which were visually similar to gadolinium-enhanced images. Furthermore, a threshold was created which could separate viable and necrotic Wilms' tumour tissue with high specificity (90%) and sensitivity (85%).

This investigation led to a publication:

H. J. Rogers, M. V. Verhagen, S. C. Shelmerdine. C. A. Clark, & P. W. Hales. An Alternative Approach to Contrast-Enhanced Imaging: Diffusion Weighted Imaging and T1-weighted Imaging Identifies and Quantifies Necrosis in Wilms Tumour. *European Radiology*. (2018). doi: 10.1007/s00330-018-5907-z.

8.1: Background

In Europe Wilms' tumour patients are treated under the SIOP approach in which they undergo pre-operative chemotherapy to reduce tumour size prior to surgery[1]. Furthermore, those with higher stage tumours and higher risk subtypes will have more intense chemotherapy prior to surgery and additional chemotherapy post-surgery (for further details please see Chapter 1.5).

Chemotherapy can cause an increase in necrotic tissue within the tumour. The volume of necrosis post-chemotherapy is informative of treatment response. It has been suggested that patients with 100% necrosis post-chemotherapy, when assessed via histological analysis, were associated with relapse free survival at 5-year follow-up[2]. Furthermore, when tumour size remains stable following treatment (without, or with minor, reduction in volume), it may appear that treatment was ineffective. However, this approach has been suggested to be unreliable, as the volume of necrosis may have increased despite small overall tumour volume changes[3]. Thus, quantifying the degree of necrosis in Wilms' tumour tissue is an important clinical indicator.

In the body, MRI can identify necrotic tissue via administration of gadolinium-based contrast agents and T_1-weighted imaging (T_1w). The contrast agent extravasates into viable tissue via the local capillary network, and as such in poorly-perfused necrotic tissue, absent or decreased enhancement may represent necrosis. This method has previously been used in Wilms' tumour[3–5] to identify necrotic regions. Furthermore, it has been shown that identifying these regions using T_1w imaging and gadolinium can assist with determining the effects of chemotherapy[6]. However, gadolinium requires venous access, raises examination costs, and can cause sides effects such as nausea and irritation at the injection site. Additionally, gadolinium is contra-indicated in those with renal failure[7], and while many Wilms' patients will not have complete renal failure, their renal function is likely to be compromised. In addition, recent reports have described gadolinium retention in neural and body tissue regardless of renal function, however currently there are no known sequelae related to this[8]. While gadolinium is still frequently administered and has many additional uses, an alternative approach to identify and quantify necrosis would be beneficial.

The apparent diffusion coefficient (ADC), derived from diffusion weighted imaging (DWI), has been shown to be related to cell density; low ADC values correlate with high cell counts in a range of paediatric body tumours[9]. ADC values have been shown to increase following chemotherapy in abdominal tumours[3] and specifically in Wilms' tumour[4]. Thus, areas of necrosis in Wilms' tumour could potentially be identified as regions with low cellular density; potentially resulting in higher ADC values. However, lower ADC values do not necessarily indicate viable tissue; necrosis by coagulation results in low ADC values which mimics high cellular density tissue[10]. However, hyper-intense regions on pre-gadolinium T_1w images can indicate areas of coagulated blood, thus a combination of ADC and pre-gadolinium T_1w may enable necrosis in Wilms' tumour to be identified and quantified, without the need for exogenous contrast agents.

Furthermore, non-Gaussian diffusion models of diffusion, such as IVIM[11], stretched exponential[12] and kurtosis[13], provide additional diffusion parameters (see Chapter 3 for details on each of these models). It has been shown that these models provide better fits (according to AIC) to DWI data compared to a mono-exponential model, when assessed in Wilms' tumour patients, both pre- and post-chemotherapy (Chapter 7).

Given that these models may provide a more accurate description of the diffusion MR signal, and provide additional information about tissue-microstructure compared to the standard mono-exponential model[14], it was hypothesised that they may also be able to identify necrotic tissue in Wilms' tumour. IVIM could be particularly beneficial in assessing

necrosis, as it accounts for the influence of blood flow on the DWI signal[15], which should be absent in necrotic tissue.

8.2: Aims and Hypotheses

It would be beneficial to non-invasively identify necrosis in Wilms' tumour; diffusion parameters and T_1w imaging may provide the means to achieve this. In this study it was hypothesised that combining non-contrast-enhanced T_1w imaging and DWI the degree of necrosis in Wilms' tumour could be estimated, as opposed to the more traditional method of using gadolinium. Different diffusion parameters were used in combination with T_1w to determine a relationship with enhancement, these models were tested by using multiple linear regressions and the preferred model was selected based on the highest adjusted R^2 value.

It was expected that an inverse relationship would be seen in majority of the parameters with an increase in ADC (mono-exponential), D and D^*(IVIM), DDC and α (stretched exponential), D_k and K (kurtosis) being significantly ($p < 0.05$) related to a decrease in mean fractional enhancement, the parameter f x D^* was also investigated and the same relationship was expected. Furthermore, a positive relationship was anticipated with f, with an increase in mean f being significantly ($p < 0.05$) related to an increase in mean fractional enhancement.

It was expected that the models with the IVIM parameters would provide the highest adjusted R^2 values.

The selected model from the regression analysis will be compared to the gadolinium method and will have little bias. A small bias will be defined as below 10% based on Bland-Altman analysis.

Additionally, this study aimed to establish a minimum threshold of enhancement, based on typical values found in necrotic tumour tissue, as defined by manual delineation of necrotic tumour regions on post-gadolinium T_1w images by two radiologists. Tissue which showed enhancement below this threshold (using either measured or predicted enhancement data) would be classified as necrotic, therefore allowing quantification of the volume of necrotic tissue in future Wilms' tumour studies.

8.3: Methods

8.3.1: Study Population

Institutional ethical approval was granted and waived the need for consent for this single centre review of clinically acquired MRI scans. A 5-year retrospective review (March 2012-2017) of the radiology imaging system was performed for all MRI abdominal studies in children with proven histological diagnosis of Wilms' tumour at Great Ormond Street Hospital (London, UK). Inclusion criteria were children who had completed a full 6-week course of chemotherapy, with MRI sequences that included both DWI with multiple b value DWI data (with a maximum b value of 1,000 s/mm^2), and T_1w sequences (pre- and post-gadolinium contrast). Cases where the post-chemotherapy size of the tumour did not cover more than 2 axial slices on diffusion imaging were excluded.

8.3.2: MRI and Contrast Agents

The details of the DWI protocol can be found in Chapter 7.3.2. Patients also received fat-supressed axial T_1-weighted imaging (T_1w) after DWI which was acquired both before and after intravenous administration of gadolinium-based contrast, using identical protocols for the pre- and post-gadolinium acquisitions: voxel size: 0.49 x 0.49 x 9.5mm, matrix size: 512 x 384 x 20, number of slices: 20. One of two T_1w sequences were performed, this was either a fast 2D spin echo sequence: TR/TE: 450ms – 550ms/8ms – 18ms, with a selective water-exciting RF pulse, or a fast low-angle shot 2D gradient echo sequence: TR/TE: 1500ms/7ms, flip angle: 15°, with a selective water-exciting RF pulse. The former sequence was acquired during free breathing in children who were anaesthetised, and the latter sequence used prospective diaphragmatic gating in those patients who were awake or moderately sedated.

Patients were administered gadolinium contrast and received the following: Dotarem 0.5mmol/ml (Manufactured by Guerbet), dosage: 0.2mls per kg body-weight. The post-contrast T_1w sequence was started 2 to 4 minutes after injection of the contrast agent. Hyoscine butylbromide (Buscopan) 20mg/ml (manufactured by Sanofi) was also administered prior to all sequences to prevent peristalsis, dosage: 0.02ml per kg body-weight, however, maximum dosage was based on patient age: 1month – 4 years = 0.25ml maximum, 4 years – 12 years = 0.5ml maximum.

8.3.4: Processing of MRI Data & ROIs

Data processing and analysis was performed using in-house routines written in Matlab R2015b (MathWorks Inc., Natick, MA, USA). All registrations were performed using NiftyReg[16] packages using affine transformations, and Regions of Interest (ROIs) were generated using Mango Software (Research Imaging Institute, UTHSCSA).

Tumour ROIs were independently drawn by two radiologists specialising in paediatric radiology (S.S: 4 years paediatric radiology; M.V: 2 years paediatric radiology). ROIs were drawn around the perimeters of each Wilms' tumour on b=0 (non-diffusion-weighted) images on each axial slice, using all clinically acquired images for guidance. The overlapping areas (between the two radiologists) were defined as the final Wilms' tumour ROIs. ROIs which displayed substantial visual mismatch between the two radiologists were reviewed until consensus was achieved. To compare similarity of size between the independently defined ROIs the Intraclass Correlation Coefficient (ICC) was calculated.

Data from each patient was processed using four different models of diffusion: mono-exponential, IVIM, stretched exponential, and kurtosis. The details of these fitting methods can be seen in Chapter 6.3.3. The resulting parameter maps were then smoothed with a 2mm Gaussian kernel using FSLmaths[17].

Fractional enhancement maps were generated from T_1w scans. All T_1w scans (pre- and post-gadolinium) were smoothed with a 2mm Gaussian kernel prior to registrations to counteract registration errors. Post-gadolinium T_1w images (post-Gd T_1w) were then registered to pre-gadolinium T_1w images (pre-Gd T_1w) using an affine registration[16]. Voxel-wise fractional enhancement maps were calculated using Equation 8.1 where the terms are measured in signal intensities.

$$\text{Fractional Enhancement} = \frac{(Post{-}Gd\ T_1w\ {-}Pre{-}Gd\ T_{1\,w})}{Pre{-}Gd\ T_{1\,w}} \quad [8.1]$$

For example, a fractional enhancement of 0.50 indicates a 50% increase in signal intensity on the T_1w image following gadolinium administration. Fractional enhancement maps were co-registered to DWI space. Additionally, pre-Gd T_1w images were normalised to mean pre-Gd T_1w signal intensity in a reference tissue in each patient, to produce semi-quantitative, normalised pre-Gd T_1w images (T_1w_{norm}). This was achieved by dividing each pre-Gd T_1w image by the mean signal intensity of an ROI placed in normal appearing

erector spinae muscles for each patient. These T_1w_{norm} images were also registered to the DWI scans.

8.3.5: Analysis and Statistics

Wilms' tumour ROIs were placed on co-registered fractional enhancement, diffusion, and T_1w_{norm} maps. Mean values were calculated for each parameter in every Wilms' tumour. The diffusion parameters included ADC, the fitted IVIM parameters (D, D^* and f), the fitted stretched exponential parameters (DDC and α), and the fitted kurtosis parameters (D_k and K). Additionally, the parameter f x D^* (from IVIM) was also investigated. Multiple linear regression was used to calculate the relationship between mean fractional enhancement (dependent variable) and a combination of the mean of a single diffusion parameter and mean T_1w_{norm} (predictor variables). Statistically significant regression coefficients were defined as having a p-value <0.05. Models were then compared based on adjusted R^2 values.

Using the selected regression model, voxel-wise predicted enhancement maps were generated for each Wilms' tumour. Using a Bland-Altman plot, whole tumour values of predicted and measured fractional enhancement were compared, to determine the similarity (confidence intervals) and level of bias (mean difference) between the two techniques. Good agreement between techniques was defined as have a level of bias $\leq 10\%$.

An upper threshold for enhancement in necrotic tissue was also generated, which would allow tumours to be separated into viable and necrotic components. Regions within each tumour which confidently represented necrosis were independently delineated by the two radiologists on the b0 images, using all clinically acquired MR sequences for guidance. The overlapping areas between the radiologists were defined as the final necrotic ROIs. These were used as the 'gold-standard' to represent the necrotic part of each tumour. The remainder of the Wilms' tumour was defined as the viable ROI. Both viable and necrotic ROIs were registered onto corresponding measured fractional enhancement maps. The fractional enhancement value of every necrotic and viable voxel was pooled across the cohort. Receiver Operator Characteristic (ROC) analysis was used to define a fractional enhancement threshold which best separated viable and necrotic tissue based on AUC (Area Under Curve) values. Tissue within a Wilms' tumour with enhancement below this threshold would be classified as necrotic, and tissue with enhancement above this threshold would be classified as having some degree of viability.

8.4: Results

8.4.1: Study Population

A total of 37 Wilms' tumours from 34 patients were included as the final cohort. The median age of patients at the time of their MRI scans was 2.6 years (*Mean:* 3.3 years, *SD*: 2.6, *Minimum:* 0.4 years, *Maximum:* 11.0 years). Patient inclusion and exclusion metrics are shown in Figure 8.1. Eight Wilms' tumour DWI acquisitions were acquired using the 7 b values range and 29 Wilms' tumour DWI acquisitions were acquired using the 8 b values range.

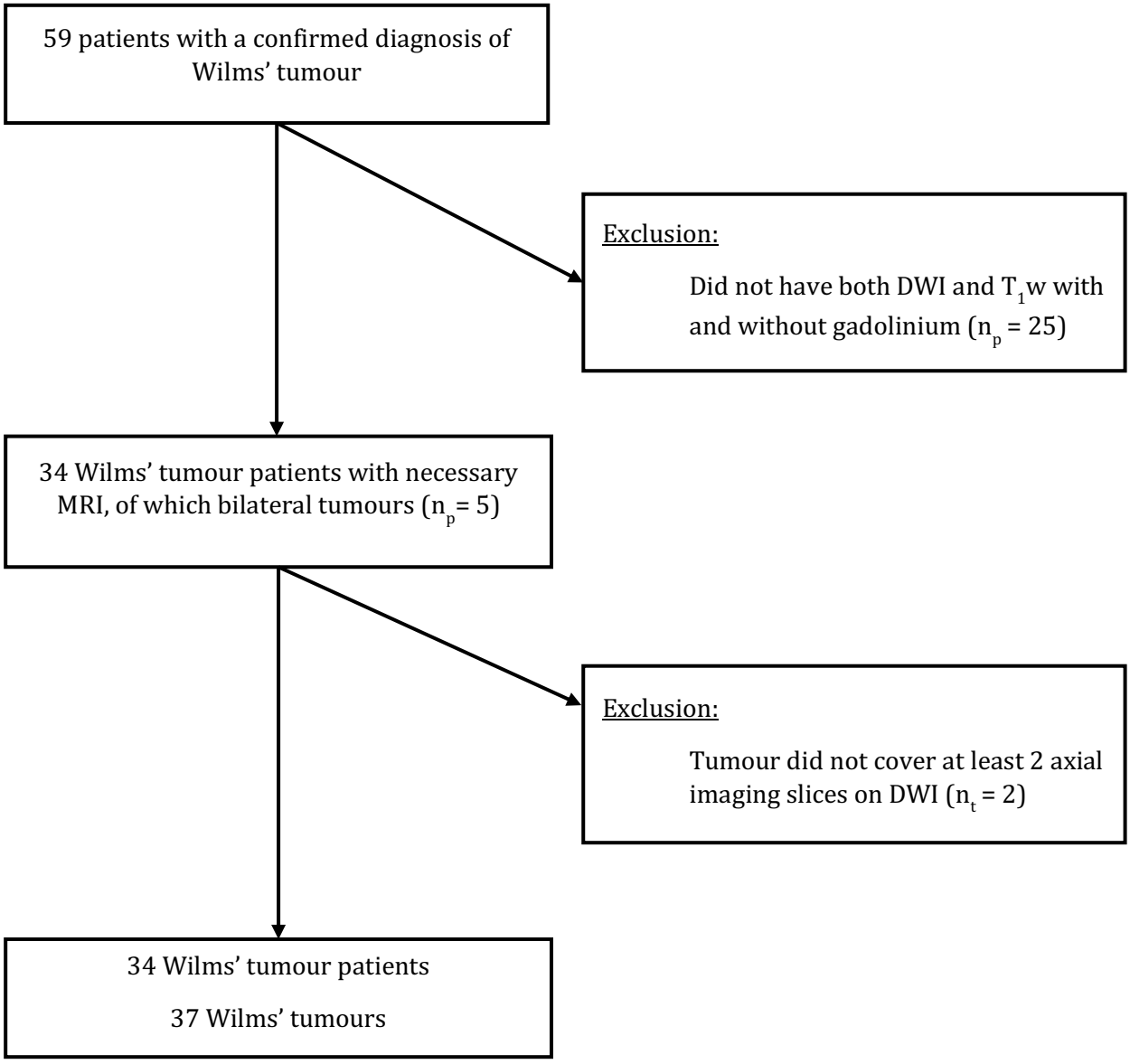

Figure 8.1. Flowchart highlighting inclusions and exclusions of Wilms' tumour data. DWI = Diffusion Weighted Imaging. T_1w = T_1 weighted imaging. n_p = number of patients, n_t = number of tumours.

8.4.2: Post-Processing

After initial delineation of the Wilms' tumour ROIs, visual inspection showed that 8/37 (21.6%) had a substantial mismatch between the radiologists; these were re-defined after consensus. The remaining 29 Wilms' tumours had a high level of agreement between radiologists with an average overlapping area of 88% (*SD:* 0.67). After adjustment of the 8 mismatched ROIs, there was high similarity in the size of the 37 Wilms' tumours as defined by the two readers, with an ICC of 0.98, (ICC prior to adjustment: 0.96). There was

also high similarity in the size of the necrotic ROIs defined by the two readers, with an ICC of 0.83.

8.4.3: Analysis

All multiple linear regression models used to predict fractional enhancement were statistically significant ($p < 0.05$), as shown in Table 8.1.

The combination of D from IVIM and T_1w_{norm} gave the most significant regression model $F(2, 34) = 13.78$, $p < 0.001$, adjusted $R^2 = 0.42$. However, this represented only a very marginal improvement compared to ADC ($F(2, 34) = 13.2$, $p < 0.001$, adjusted $R^2 = 0.40$). While the other 7 models all reached significance ($p < 0.05$), the higher p values and comparatively low adjusted R^2 values indicated that they did not predict fractional enhancement levels as well. Due to the similarity in performance between the regression models based on D (IVIM) and ADC (mono-exponential), and the fact that ADC data are more widely acquired clinically (as fewer b-values are required), we chose to focus on the ADC-based model for further analysis.

Table 8.1. The *p*-values, adjusted R^2 and correlation coefficients (β) of the nine multiple regression models used to predict Fractional Enhancement, based on a combination of T_1w_{norm} and one of the diffusion parameters. ADC, *D*, *D**, *f* x *D**, *DDC*, D_k were all measured in standard units of mm/s^2. *f*, α, *K*, and T_1w_{norm} are unitless

Diffusion Parameter	β_0 (Intercept)	β_1 (Diffusion)	β_2 (T_1w_{norm})	Model *p* Value	Model Adjusted R^2
ADC (mono-exponential)	1.85	-408.4	-0.4	5.7×10^{-5}	0.40
D (IVIM)	1.83	-419.64	-0.4	4.2×10^{-5}	0.42
*D** (IVIM)	1.18	-1.09	-0.3	0.017	0.17
f (IVIM)	1.01	1.07	-0.34	0.025	0.15
f x *D** (IVIM)	1.2	-5.53	-0.33	0.023	0.15
DDC (stretched exponential)	1.11	-1958.39	-0.27	0.003	0.25
α (stretched exponential)	3.04	-2.28	-0.32	0.001	0.29
D_k (kurtosis)	1.91	-324.16	-0.41	0.001	0.31
K (kurtosis)	1.17	3.32×10^{-7}	-0.35	0.032	0.14

Note. ADC = apparent diffusion coefficient. IVIM = intravoxel incoherent motion. *D* = thermally-driven, 'slow' diffusion, *D** = flow-driven, 'fast' diffusion, *f* = volume fraction associated with 'fast' diffusion, *DDC* = distributed diffusion coefficient, α = stretching parameter, D_k = diffusion kurtosis coefficient, *K* = kurtosis

Figure 8.2 demonstrates the relationship between both ADC and T_1w_{norm} with fractional enhancement. Both ADC ($p < 0.001$) and T_1w_{norm} ($p = 0.001$) added significantly to the prediction, with both increased ADC and increased T_1w_{norm} being associated with reduced fractional enhancement. The standard error of the estimate was 0.24.

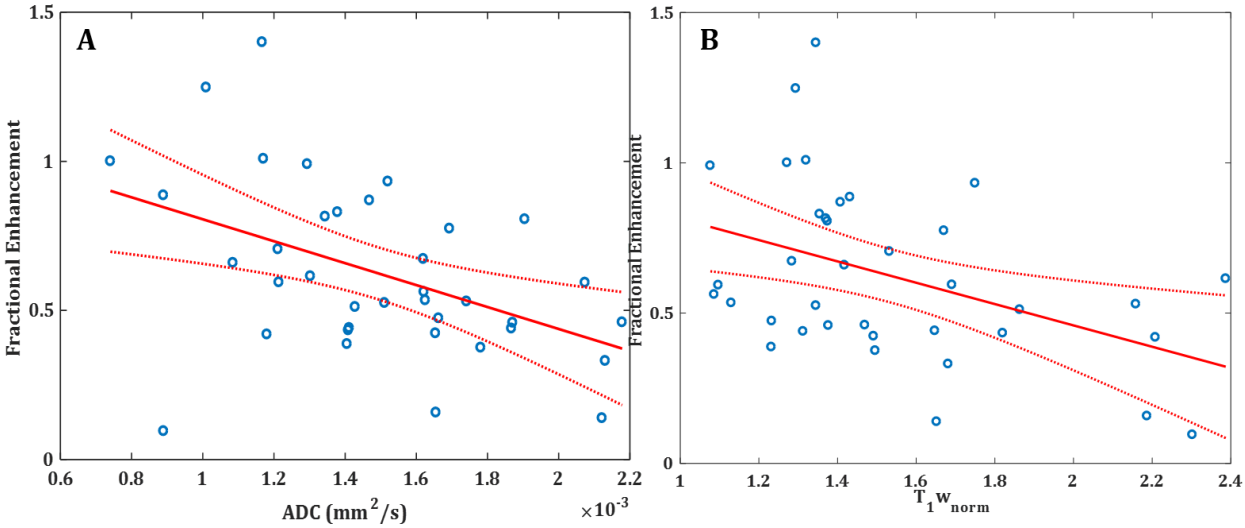

Figure 8.2. (A) Linear regression of mean ADC (apparent diffusion coefficient) versus mean fractional enhancement in 37 Wilms' Tumours, adjusted $R^2 = 0.19$ **(B)** Linear regression of mean T_1w_{norm} (normalised semi-quantitative T_1weighted imaging) versus mean fractional enhancement in 37 Wilms' Tumours, adjusted $R^2 = 0.16$. For the multiple linear regression model (with both ADC and T_1w_{norm} as predictors), the adjusted R^2 was 0.40 ($p < 0.001$).

Using the ADC- T_1w_{norm} model, predicted enhancement was calculated according to the regression model given by Equation 8.2 below, which was derived using the 'fitlm' algorithm in Matlab:

$$\text{Predicted Enhancement} = 1.85 - (408.4 * ADC) - (0.4 * T_1w_{norm}) \quad [8.2]$$

where ADC is measured in mm/s^2.

Comparisons between fractional enhancement and predicted enhancement maps in three representative patients are illustrated in Figure 8.3. Both highlight similar regions of enhancing and non-enhancing tissue.

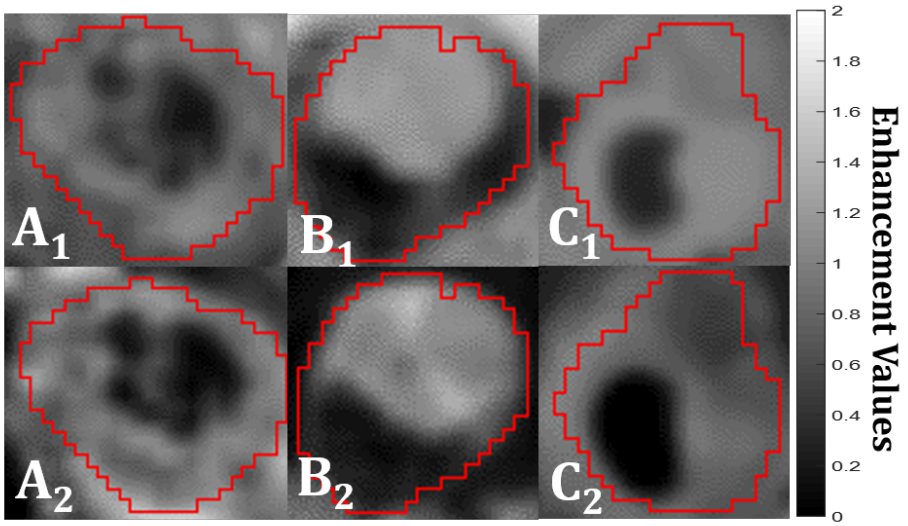

Figure 8.3. Examples of single axial slices from three representative Wilms' tumours. **(A_1, B_1, C_1)** Fractional enhancement maps of the Wilms' tumours (outlined in red), measured using gadolinium. **(A_2, B_2, C_2)** The same slices of the same Wilms' tumours from predicted enhancement maps, predicted using Equation 8.2 (without gadolinium). Increased signal represents greater enhancement, and hence more viable tissue. Tumour details: A – subtype: mixed, age at scan: 11 years, B – subtype: blastemal, age at scan: 1.8 years, C – subtype: mixed, age at scan: 1.08 years.

The level of agreement between fractional enhancement and predicted enhancement is illustrated in the Bland-Altman plot in Figure 8.4. There was a small bias (9.3%) in predicted values overestimating the level of enhancement across a wide range of enhancement levels (Mean difference (measured FE – predicted FE) = -0.093, 95% CI = [-0.52, 0.34]).

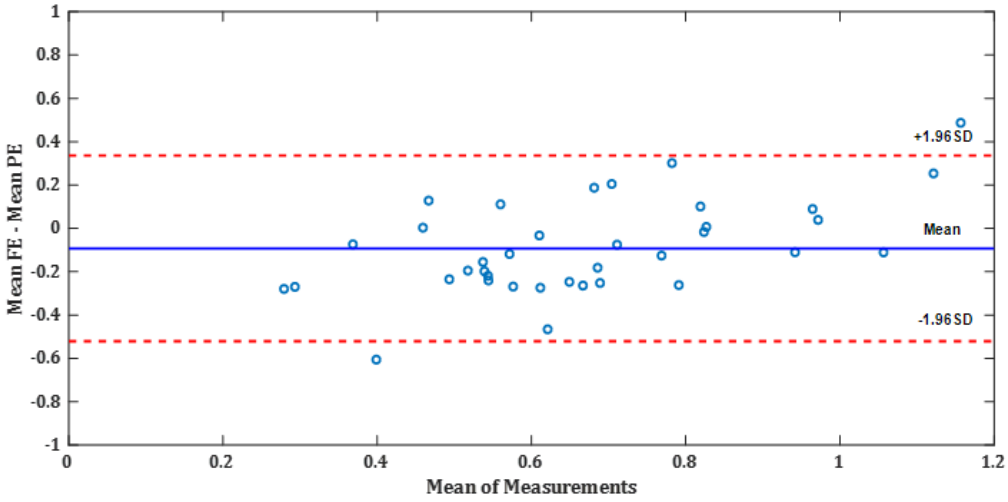

Figure 8.4. A Bland-Altman plot showing the level of agreement in mean enhancement values in 37 Wilms' tumours, as calculated using fractional enhancement (FE) and predicted enhancement (PE) from the ADC-T_1w_{norm} model (*Equation 8.2*).

ROC analysis provided an optimal threshold to distinguish between viable and necrotic tissue, based on fractional enhancement (Figure 8.5a). The upper threshold was 0.33 (i.e. voxels showing less than 33% signal enhancement on T_1w imaging, after administration of gadolinium, were classified as necrotic). This threshold provided a sensitivity of 85% and specificity of 90% for identifying the 'gold-standard' necrotic tissue (i.e. all voxels contained within the necrotic ROIs defined by the two radiologist), with an AUC of 0.93.

Figure 8.5b displays a box and whisker plot of the fractional enhancement values in all voxels within the manually defined necrotic and viable ROIs across the entire cohort. An independent sample t-test revealed a significant difference between fractional enhancement values in the viable (*Mean:* 0.73 *SD:* 0.33) and necrotic (*Mean:* 0.14 *SD:* 0.2) voxels; $t(195364) = -446.96$, $p < 0.001$. The optimum threshold (0.33) which separates necrotic and viable tumour tissue is also highlighted in Figure 8.5b.

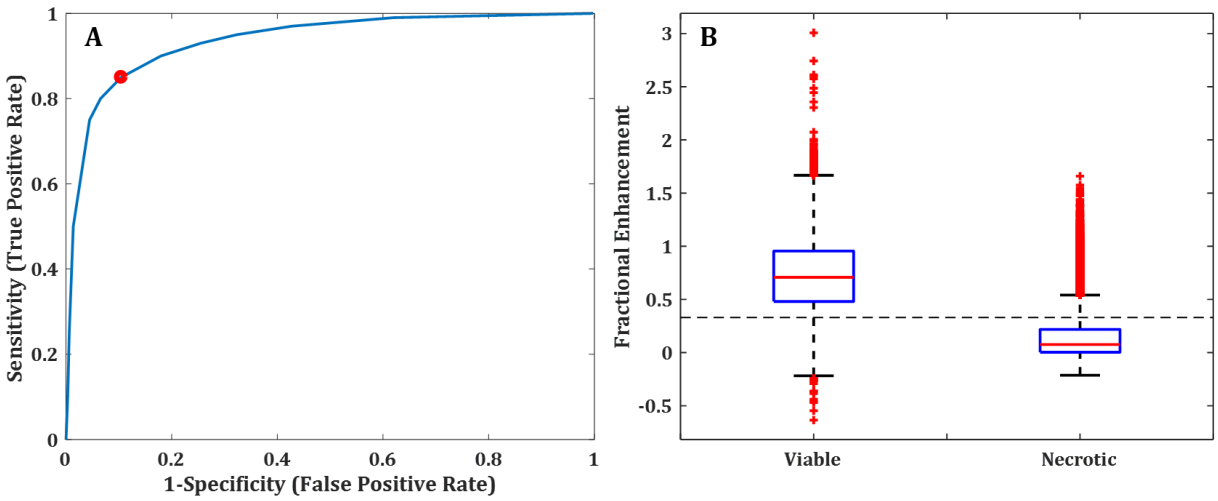

Figure 8.5. (A): Receiver operator characteristic curve to determine a threshold which best separates necrotic and viable Wilms' tumour tissue. The optimum upper threshold (0.33), whereby voxels displaying enhancement above this value are classified as viable, is highlighted in red. For this threshold the area under the curve was 0.93, sensitivity was 85%, and the specificity was 90%. **(B):** Box and whisker plot displaying fractional enhancement of every voxel from the 37 Wilms' tumours which were either classified as necrotic or viable. The dotted line reflects the optimum threshold (0.33 fractional enhancement) for this separation based on ROC analysis which is shown in **5A**.

8.5: Discussion

This study investigated whether necrosis could be identified without using gadolinium contrast-enhanced T_1w images in Wilms' tumour. Good agreement (defined as ≤10%) was found between mean tumour enhancement values calculated using non-gadolinium-based metrics (ADC and T_1w_{norm}), and the level of enhancement measured using gadolinium in the same tumours. Additionally, a threshold of maximum enhancement in necrotic tissue was determined, which separated viable and necrotic tissue in good agreement with manually delineated necrotic tumour regions, as defined by two specialist paediatric radiologists. As such, this threshold could be used to quantify the total fraction of necrotic tissue in Wilms' tumours in future studies, using either measured or predicted enhancement values.

Necrosis within Wilms' tumours can indicate chemotherapy response, with high volumes of necrosis representative of 'good response'[18]. Quantifying the percentage of necrosis in Wilms' tumour has previously been challenging as histological methods usually only sample a sub-section of tissue. Thus, measuring the level of necrosis of the entire tumour volume using imaging-based assessment, particularly without exogenous contrast, is greatly beneficial. In addition, although 'contrast enhancement' is currently used in a qualitative manner to identify viable/necrotic tumour tissue, this is rarely done in a quantitative manner, and as such the intra-tumour volume fraction of necrotic tissue is rarely reported radiologically. By deriving a threshold value for the minimum level of enhancement necessary for tissue to be classified as viable (which could be measured using either gadolinium or the non-gadolinium-based model presented here), we hope that the degree of necrosis in a tumour can be quantified in future studies. Additionally, in instances of bilateral Wilms' tumour whole tumour resection is not possible and thus total necrosis fractions cannot be quantified using histological analysis. Furthermore, DWI and T_1w imaging are routinely acquired in Wilms' tumour patients, so no additional scan time is needed, aiding the transference to clinical practice.

A further benefit of identifying this necrotic tissue is that it can assist with identifying subtypes. Previous work has shown that some Wilms' tumour subtypes could be separated into lower and higher-risk groups using ADC[4]. This work highlighted that using lower quartile ADC values provided better separation compared to mean ADC values, thus, removing higher diffusion values (which are likely to represent necrotic

tissue) improved subtype identification. Therefore, providing another use for the proposed method – this is discussed in more detail in Chapter 9.

The current gadolinium-based method of identifying non-enhancing tissue has some possible limitations due to potential adverse reactions in patients (including nausea, headaches and irritation), and the as-yet unknown impact of the accumulation of contrast agent in patients undergoing repeated follow-up imaging[8]. Additionally, gadolinium may not always be appropriate in Wilms' tumour, for example if the patient has renal failure; which will be dependent on co-morbid disorders, tumour staging, treatment timeline and whether the tumour is bilateral[19,20]; additionally, chemotherapy drugs can lead to nephrotoxicity[21]. For these reasons, alternative approaches for predicting enhancement and identifying and quantifying necrotic tissue without gadolinium, are potentially beneficial.

ADC is a well-defined diffusion parameter; however, multiple b value DWI data allows non-Gaussian diffusion models to be applied, which provide additional fitted parameters. This study compared ADC and non-Gaussian diffusion parameters in predicting fractional enhancement. Out of all the models considered here, IVIM had the highest adjusted R^2 value. Previous work has shown that D and f have higher accuracy in distinguishing between enhancing and non-enhancing kidney lesions compared to ADC[22]. Furthermore, bi-exponential models have been suggested rather than mono-exponential for more reliable diffusion estimates of healthy kidney tissue[23]. Interestingly, regression models which incorporated the f and f x D^* parameters did not reach high significance for predicting enhancement. f is assumed to represent the contribution to the DWI signal due to blood flowing in the randomly orientated capillary network[11], and f x D^* represents a surrogate measure of blood flow[24]. Due to the lack of blood flow in necrotic tissue it would be expected that these parameters would be better predictors of fractional enhancement, however our results suggest this is not the case. This may be because it is beyond the sensitivity of the IVIM model to identify the small level of perfusion in viable Wilms' tumours compared to non-enhancing tissue.

D produced the strongest regression model, however, the difference between the predictive power of D and ADC was minimal. Due to the similarity in the performance of these two predictors, and the fact that ADC values are routinely acquired in clinical practice (whereas D requires longer, multiple b value acquisitions), ADC represents the preferred option for predicting fractional enhancement when combined with T_1w_{norm}. This combination is needed as ADC alone cannot account for necrosis via coagulation, and as

can be seen (Figure 8.2) the regression is much stronger when T_1w_{norm} is added as a predictor.

This study had several limitations. Firstly slightly different b values were used for a small number of our patients (8/37 were acquired with 7 b values and 29/37 were acquired with 8 b values), however previous work has shown high reproducibility between ADC values acquired on different scanners with varying b values[25]. Furthermore, comparisons of ADC when acquired with either 7 or 8 b values (using the exact same MRI protocols as in the present study) in healthy renal tissue showed a small coefficient of variation (6.27%), and almost no bias when assessed via Bland-Altman analysis (<1%) (Chapter 6). Therefore, the b value alterations would be unlikely to influence our analysis.

Secondly, when comparing the measured and predicted mean enhancement values, the predicted values were slightly overestimated. However, this bias was small (9%) and may be due to registration errors between T_1w and ADC maps.

Thirdly, our sample size was fairly small, and a more robust model may be possible with a larger cohort. Furthermore, we did not assess tumour necrosis independently using histological methods. However, it is important to note that for histological analysis of Wilms' tumour, only a sub-section of the tumour is sampled, and this may not accurately reflect the total necrosis volume. As such, we preferred in this study to use visual assessment of the entire tumour volume, using all clinically available MRI scans, to ensure the entire tumour volume was assessed. For this method we must rely on the ROIs defined by the two radiologists, thus if they were incorrect then the tumour and necrosis selection would be wrong. This in turn would impact the mean values of the diffusion parameters and T_1w_{norm} and therefore generate incorrect regression models and the threshold. However, as the radiologists had many years of experience in this paediatric population, and both selected similar regions when they created the ROIs independently we can place confidence in the selected regions.

An additional limitation may arise from the possible reliance of our model on the specifics of the T_1w protocol used in this study. Alterations in the delay time between gadolinium administration and T_1w imaging may lead to different levels of enhancement on contrast-enhanced T_1w scans. However, as our model uses the fractional difference between the pre- and post-contrast T_1w signal, provided the T_1w protocol remains consistent between these two acquisitions, the influence of variations in the specifics of the T_1w acquisition between different institutions should mostly cancel out.

Finally, gadolinium is frequently administered for indications broader than necrosis assessment, for example, vascular anatomy, and detecting lesions in a variety of organs. We acknowledge that the proposed method cannot entirely replace gadolinium. Despite this, our method would be a suitable alternative for those with severe renal impairment and limit the cumulative gadolinium exposure for patients who have repeated follow-up MRI scans. Gadolinium-free MRI examinations are currently being investigated for paediatric oncology[26] and the proposed model could facilitate this, given that it uses data acquired as part of the clinical standard.

8.6: Conclusions and Summary

Volume of necrotic tissue in Wilms' tumour is informative of treatment response. Currently visualisation of necrotic regions relies on gadolinium-contrast enhanced imaging. This study sought to determine whether an alternative method of combining DWI and T_1w could identify and quantify necrosis, as there are associated risks and complications with gadolinium.

All diffusion parameters were significantly related to enhancement, and the IVIM parameter, *D*, did produce the strongest regression model (determined by adjusted R^2) as predicted by the hypotheses. However, *D**, *f*, and *f x D** models did not produce as strong regressions as anticipated.

Although the *D*-T_1w_{norm} model produced the highest adjusted R^2, it was not much greater than ADC, and given the wider clinical use of ADC, the ADC-T_1w_{norm} was chosen. This model generated maps which were visually similar to the gadolinium method and also had good agreement with the gadolinium data ($\leq 10\%$).

A threshold of maximum enhancement in necrotic regions was generated based on the chosen model (enhancement < 0.33 = necrotic), allowing the percentage of necrotic tissue to be quantified in future Wilms' tumour studies using imaging-based methods without the need for an exogenous contrast agent.

Main Finding: ADC and T_1w imaging can provide an alternative approach to identify and quantify necrosis in Wilms' tumour without the need for contrast agents.

References: Chapter 8

1. Heuvel-Eibrink, M. M. van den *et al.* Position paper: Rationale for the treatment of Wilms tumour in the UMBRELLA SIOP–RTSG 2016 protocol. *Nat. Rev. Urol.* **14**, 743–752 (2017).
2. Boccon-Gibod, L. *et al.* Complete necrosis induced by preoperative chemotherapy in Wilms tumor as an indicator of low risk: report of the international society of paediatric oncology (SIOP) nephroblastoma trial and study 9. *Med. Pediatr. Oncol.* **34**, 183–190 (2000).
3. McDonald, K., Sebire, N. J., Anderson, J. & Olsen, O. E. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy-feasibility study. *Pediatr. Radiol.* **41**, 99–106 (2011).
4. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
5. Littooij, A. S., Humphries, P. D. & Olsen, Ø. E. Intra- and interobserver variability of whole-tumour apparent diffusion coefficient measurements in nephroblastoma: a pilot study. *Pediatr. Radiol.* **45**, 1651–1660 (2015).
6. Littooij, A. S. *et al.* Apparent diffusion coefficient as it relates to histopathology findings in post-chemotherapy nephroblastoma: a feasibility study. *Pediatr. Radiol.* **47**, 1608–1614 (2017).
7. Perazella, M. A. Gadolinium-contrast toxicity in patients with kidney disease: nephrotoxicity and nephrogenic systemic fibrosis. *Curr. Drug Saf.* **3**, 67–75 (2008).
8. McDonald, R. J. *et al.* Intracranial Gadolinium Deposition after Contrast-enhanced MR Imaging. *Radiology* **275**, 772–782 (2015).
9. Humphries, P. D., Sebire, N. J., Siegel, M. J. & Olsen, Ø. E. Tumors in Pediatric Patients at Diffusion-weighted MR Imaging: Apparent Diffusion Coefficient and Tumor Cellularity. *Radiology* **245**, 848–854 (2007).
10. LaViolette, P. S. *et al.* Precise ex vivo histological validation of heightened cellularity and diffusion-restricted necrosis in regions of dark apparent diffusion coefficient in 7 cases of high-grade glioma. *Neuro-Oncol.* **16**, 1599–1606 (2014).
11. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).

12. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
13. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
14. Yuan, J. *et al.* Non-Gaussian Analysis of Diffusion Weighted Imaging in Head and Neck at 3T: A Pilot Study in Patients with Nasopharyngeal Carcinoma. *PLOS ONE* **9**, e87024 (2014).
15. Iima, M. & Le Bihan, D. Clinical Intravoxel Incoherent Motion and Diffusion MR Imaging: Past, Present, and Future. *Radiology* **278**, 13–32 (2015).
16. Modat, M. *et al.* Global image registration using a symmetric block-matching approach. *J. Med. Imaging* **1**, (2014).
17. Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W. & Smith, S. M. FSL. *NeuroImage* **62**, 782–790 (2012).
18. Godzinski, J. The current status of treatment of Wilms' tumor as per the SIOP trials. *J. Indian Assoc. Pediatr. Surg.* **20**, 16–20 (2015).
19. Breslow, N. E. *et al.* END STAGE RENAL DISEASE IN PATIENTS WITH WILMS TUMOR: RESULTS FROM THE NATIONAL WILMS TUMOR STUDY GROUP AND THE UNITED STATES RENAL DATA SYSTEM. *J. Urol.* **174**, 1972–1975 (2005).
20. Grigoriev, Y. *et al.* Treatments and outcomes for end-stage renal disease following Wilms tumor. *Pediatr. Nephrol.* **27**, 1325–1333 (2012).
21. Perazella, M. A. Onco-Nephrology: Renal Toxicities of Chemotherapeutic Agents. *Clin. J. Am. Soc. Nephrol.* **7**, 1713–1721 (2012).
22. Chandarana, H., Lee, V. S., Hecht, E., Taouli, B. & Sigmund, E. E. Comparison of Biexponential and Monoexponential Model of Diffusion Weighted Imaging in Evaluation of Renal Lesions: Preliminary Experience. *Invest. Radiol.* **46**, 285 (2011).
23. Zhang, J. L. *et al.* Variability of renal apparent diffusion coefficients: limitations of the monoexponential model for diffusion quantification. *Radiology* **254**, 783–792 (2010).
24. Le Bihan, D. & Turner, R. The capillary network: a link between IVIM and classical perfusion. *Magn. Reson. Med.* **27**, 171–178 (1992).
25. Grech-Sollars, M. *et al.* Multi-centre reproducibility of diffusion MRI parameters for clinical sequences in the brain. *Nmr Biomed.* **28**, 468–485 (2015).

26. Muehe, A. M. *et al.* How to Provide Gadolinium-Free PET/MR Cancer Staging of Children and Young Adults in Less than 1 h: the Stanford Approach. *Mol. Imaging Biol.* 1–12 (2017). doi:10.1007/s11307-017-1105-7

Chapter 9: Identification of Wilms' Tumour Subtypes using Non-Gaussian Models of Diffusion

This chapter describes an original research study which aimed to identify Wilms' tumour subtypes (blastemal, epithelial, mixed, stromal, regressive) using diffusion parameters. Parameters from four models of diffusion were compared to determine whether they could separate Wilms' tumour subtypes based on one-way ANOVAs. Whole tumour values were compared to viable tumour regions to determine whether limiting analysis to viable tumour tissue increases the amount of subtype separations. It was found that all parameters other than *f* could separate some subtypes, and that limiting analysis to viable tumour tissue improved these identifications. *D** and *K* provided the most separations and were able to distinguish blastemal from epithelial subtypes, which has previously not been possible.

9.1: Background

The SIOP-2001[1] approach categorises Wilms' tumours into distinct histological subtypes based on the predominant viable cell type which remains following chemotherapy. The main subtypes are completely necrotic (low risk), stromal, epithelial, mixed (intermediate risk) and blastemal (high risk). There are differences in outcome and treatment between risk groups, with the higher risk subtype having the poorest outcome and benefitting from more intense treatment. Further details describing how these subtypes are identified, treated and their outcomes are given in Chapter 1.5.

Currently subtypes can only be identified via histological analysis post-surgery, and thus it would be beneficial to be able to identify these subtypes in a non-invasive manor, at an earlier stage, in order to guide the severity of pre-operative treatment or surgical resection. Given that ADC is related to cell density[2], and subtypes will have distinct cellular compositions, this parameter may help in identifying subtypes. However, using the mean or median ADC value of the entire tumour may not provide accurate measurements. Areas of necrotic tissue could corrupt average, whole-tumour values; as necrotic tissue would be likely to have higher ADC values than viable tumour tissue[2,3]. Furthermore, it has been shown that in post-chemotherapy Wilms' tumour tissue, median ADC values differed depending on whether or not necrotic tissue was excluded[4].

Previous work has shown that when avoiding areas of high ADC (thought to be necrotic tissue), some subtypes could be distinguished from one another[3]. Post-chemotherapy ADC maps were used for analysis as this was the closest time point to histology, and it was found that the higher risk blastemal subtype had significantly lower ADC values compared to the intermediate subtypes; stromal, regressive and mixed. However, no differences in ADC were found between epithelial and blastemal subtypes, and unfortunately these subtypes confer a different diagnosis. In addition, another study showed that within individual Wilms' tumours the proportion of stromal tissue was correlated with ADC[4]. Further details and more in-depth discussion of both experiments are given in Chapter 4.2.4.

Non-Gaussian models of diffusion have not been explored in subtype separation in Wilms' tumour. It was shown, however, that post-chemotherapy Wilms' tumours were better described by non-Gaussian diffusion models compared to the mono-exponential model according to AIC (Chapter 7). Out of 62 post-chemotherapy Wilms' tumours all but one favoured non-Gaussian models (stretched exponential[5], IVIM[6] and kurtosis[7]). As Wilms' tumour tissue was better described by these models and additional diffusion parameters are provided by these models, perhaps they may also assist with subtype separation.

9.2: Aims and Hypotheses

This experiment builds on previous work by utilising the threshold generated in Chapter 8, which identified necrotic tissue in Wilms' tumours using ADC and T_1w. After excluding necrotic tissue, four models of diffusion (mono-exponential, IVIM, stretched exponential and kurtosis) were used to determine whether any of the diffusion parameters could distinguish between Wilms' tumour subtypes.

It was expected that significant differences ($p < 0.05$) would be seen between the subtypes based on the diffusion values. Given that the non-Gaussian models provided better fits to the diffusion data it was expected that these parameters may provide more distinctions between subtypes compared to ADC.

Based on past research, it was also expected that the blastemal and epithelial subtypes will have lower ADC values compared to the other subtypes and that stromal will have the highest values.

As an additional aim, whole tumour values (including necrotic tissue) were compared to the viable ROIs to establish whether excluding necrotic regions improved the subtype separations. It was hypothesised that more distinctions would be found by limiting the ROIs to viable tumour tissue.

9.3: Methods

9.3.1: Study Population

Patients were identified based on methods described in Chapter 7.3.1. Inclusion criteria were children who had completed a full 6-week course of chemotherapy, with MRI sequences that included both DWI and T_1w (without contrast) sequences. Cases where the post-chemotherapy size of the tumour did not cover more than 2 axial slices on diffusion imaging were excluded. Tumours were also excluded if they did not have a histologically confirmed subtype according to SIOP-2001 protocol[1].

9.3.2: MRI

Diffusion and T_1w imaging (without gadolinium-contrast) were acquired according to the methods described in Chapter 8.3.2. DWI data was acquired on the 1.5T Siemens Avanto scanner and both the 7 and 8 b value ranges were included. All patients underwent DWI followed by T_1w imaging.

9.3.3: Processing of MRI Data & ROIs

Data from each patient was processed using four different models of diffusion: mono-exponential, IVIM[6], stretched exponential[5], and kurtosis[7]. The details of these fitting methods can be seen in Chapter 6.3.3.

Whole tumour volume ROIs were generated according to the methods described in Chapter 8.3.4.

9.3.4: Analysis

A flowchart detailing the processing and analysis of the data is shown in Figure 9.1.

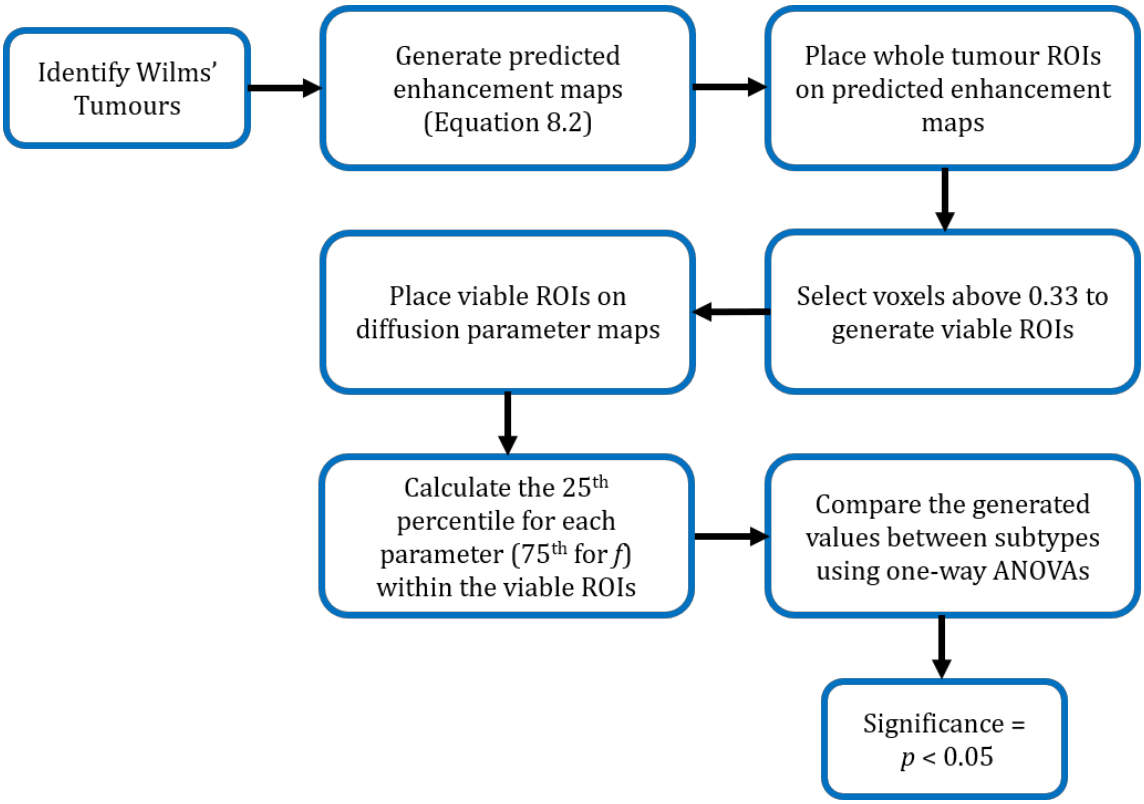

Figure 9.1. Flowchart detailing the steps involved to analyse the Wilms' tumour subtypes

To determine whether Wilms' tumour subtypes could be distinguished from one another based on diffusion measurements, and to test the clinical utility of applying the predicted enhancement model described in Chapter 8, voxel-wise predicted enhancement maps were generated according to equation 8.2 (Chapter 8). Whole tumour ROIs were placed on the predicted enhancement maps and voxels which had enhancement values under 0.33 were identified; these voxels represented areas of necrosis (see Chapter 8). New viable tumour ROIs were then generated which excluded these regions, so that analysis would only focus on the viable portions of the tumours.

Viable ROIs were then placed on the diffusion parameter maps: ADC (mono-exponential) D, D^*, f (IVIM), DDC, α (stretched exponential), D_k, K (kurtosis), additionally D^* x f parameter maps were also explored. For each tumour and parameter (except f) the 25th percentile was calculated within the ROIs. This was shown previously to best stratify Wilms' tumour subtypes based on the values in the most cellular 'subpopulation' of Wilms'

tumour tissue[3] and should further reduce the chance of including the unwanted necrotic tissue, which would be likely to have higher diffusion values due to lower cell counts[2]. For the *f* parameter the 75th percentile was calculated for each tumour instead of the 25th, as it has been suggested that lower values of *f* are related to necrosis[8].

Statistical analyses were performed using IBM SPSS (Version 24). The calculated parameter values were then compared between subtypes using a one-way ANOVA. For this statistical test certain assumptions must be met; each subtype must not have any significant outliers (defined as 3 x the interquartile range), the dependent variable (25th or 75th percentile) should be approximately normally distributed for each category of the independent variable (subtype), and the data must not violate the homogeneity of variances. To test for normality of data a Shapiro-Wilk test was performed. As one-way ANOVAs are considered robust against the normality assumption, the data only needed to be approximately normally distributed, for this reason $p > 0.01$ was selected as the criterion for data to be considered normally distributed. Levene's test for homogeneity of variances was used, where $p \geq 0.05$ would indicate no violation. If the data passed all of the assumptions then a one-way ANOVA was performed, where a significant result was defined as $p < 0.05$. Post-hoc Tukey tests were used to determine which subtypes differed from each other, where $p < 0.05$ indicated a significant difference. If the data violated the homogeneity of variances, then a Welch test was used followed by a Games-Howell post hoc test (again with $p < 0.05$ indicting a significant result).

In addition, whole tumour volume ROIs were also used to determine whether viable ROIs provided any improvement in subtype identification, using the same methodology as described above (excluding the removal of necrotic regions from the ROI).

9.4: Results

9.4.1: Study Population

A total of 48 Wilms' tumours were included in the final cohort; exclusions and inclusions can be seen in Figure 9.2. The mean age of patients at the time of their MRI scans were 3.47 years (*SD:* 2.91, median: 2.66 years, minimum: 0.44 years, maximum: 12.17 years). The subtypes of the tumours were as follows: blastemal: 7, epithelial: 7, stromal: 11, mixed: 15, regressive:7, completely necrotic: 1. As there was only one tumour classed as completely necrotic it was excluded from all analyses.

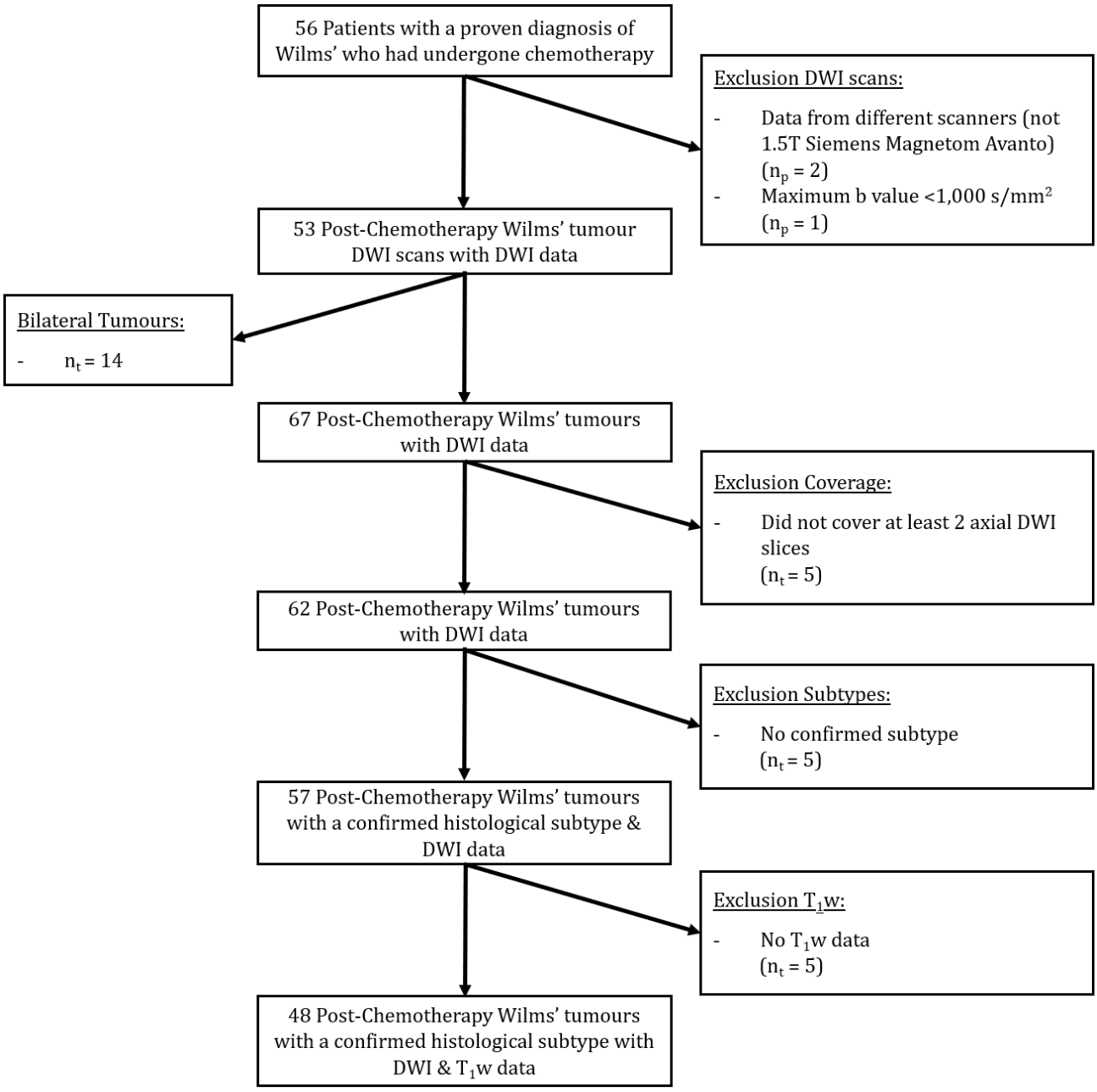

Figure 9.2. Flowchart of study population showing exclusion criteria. DWI = diffusion weighted imaging. T_1w = T_1 weighted imaging. n_p = number of patients, n_t = number of tumours.

9.4.2: Wilms' Tumour Subtypes – Descriptive Statistics

Using the threshold of 0.33 enhancement, the percent of viable tumour tissue per tumour was calculated. The average percent of viable tumour tissue was 94% (*SD:* 11.6, minimum: 49.4%, maximum: 100%).

The mean, and 25th percentile (75th percentile for f), were averaged over each subtype group per parameter for the viable and whole tumour ROIs. Tables 9.1 and 9.2 detail these results. The mean values of *DDC* were very high, which was likely due to the fitting method used to generate this parameter, this is discussed in more detail in Chapter 6.5.

Table 9.1. Descriptive statistics averaged over each subtype for each parameter. Values were obtained from viable ROIs. The 25th percentile is displayed for every parameter other than *f* where the 75th percentile is given. ADC, *D*, *D**, *f* x *D**, *DDC* and D_k are all measured in standard units of mm/s^2. *f*, α, and *K* are unitless

	Blastemal (n = 7)		**Epithelial (n = 7)**		**Stromal (n = 11)**		**Mixed (n = 15)**		**Regressive (n = 7)**	
	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile
ADC (mono-exponential)	1.24×10^{-3} (2.06×10^{-4})	0.92×10^{-3}	1.02×10^{-3} (2.11×10^{-4})	0.80×10^{-3}	1.70×10^{-3} (2.47×10^{-4})	1.50×10^{-3}	1.39×10^{-3} (2.83×10^{-4})	1.10×10^{-3}	1.41×10^{-3} (3.5×10^{-4})	1.10×10^{-3}
D (IVIM)	1.14×10^{-3} (2.12×10^{-4})	0.83×10^{-3}	0.92×10^{-3} (1.93×10^{-4})	0.73×10^{-3}	1.60×10^{-3} (2.32×10^{-4})	1.39×10^{-3}	1.30×10^{-3} (2.74×10^{-4})	1.01×10^{-3}	1.33×10^{-3} (3.43×10^{-4})	1.02×10^{-3}
*D** (IVIM)	90.3×10^{-3} (11.0×10^{-2})	4.09×10^{-3}	69.2×10^{-3} (7.72×10^{-2})	6.71×10^{-3}	66.6×10^{-3} (14.5×10^{-2})	3.34×10^{-3}	43.3×10^{-3} (6.52×10^{-2})	3.97×10^{-3}	88.2×10^{-3} (5.16×10^{-2})	2.75×10^{-3}
f (IVIM)	0.18 (0.04)	0.11	0.15 (0.03)	0.09	0.18 (0.06)	0.10	0.15 (0.04)	0.09	0.71 (0.02)	0.75
f x *D** (IVIM)	1.4×10^{-2} (1.92×10^{-2})	6×10^{-4}	1.07×10^{-2} (1.18×10^{-2})	7.5×10^{-4}	1.29×10^{-2} (3.47×10^{-2})	4.7×10^{-4}	0.71×10^{-2} (1.92×10^{-2})	4.6×10^{-4}	1.35×10^{-2} (0.92×10^{-2})	4×10^{-4}

DDC (stretched exponential)	0.56 (1.47)	0.93×10^{-3}	8.79 (23.2)	0.79×10^{-3}	1.27×10^{-2} (382.72)	1.62×10^{-3}	0.14 (0.49)	1.14×10^{-3}	$5.18 \times 10^{+6}$ ($1.37 \times 10^{+7}$)	1.14×10^{-3}
α (stretched exponential)	0.81 (3.71×10^{-2})	0.74	0.76 (4.08×10^{-2})	0.70	0.87 (4.28×10^{-2})	0.83	0.85 (3.86×10^{-2})	0.79	0.87 (5.41×10^{-2})	0.81
D_k (kurtosis)	1.85×10^{-3} (2.99×10^{-4})	1.32×10^{-3}	1.89×10^{-3} (4.47×10^{-4})	1.21×10^{-3}	2.33×10^{-3} (4.28×10^{-4})	1.87×10^{-3}	1.99×10^{-3} (3.53×10^{-4})	1.45×10^{-3}	2.01×10^{-3} (4.53×10^{-4})	1.37×10^{-3}
K (kurtosis)	1.2 (0.39)	0.39	1.36 (0.32)	1.09	0.58 (0.14)	0.34	0.89 (0.34)	0.55	0.74 (0.29)	0.30

Table 9.2. Descriptive statistics averaged over each subtype for each parameter. Values were obtained from whole tumour ROIs. The 25th percentile is displayed for every parameter other than *f* where the 75th percentile is given. ADC, *D*, *D**, *f* x *D**, *DDC* and D_k are all measured in standard units of mm/s^2. *f*, α, and *K* are unitless

	Blastemal (n = 7)		**Epithelial (n = 7)**		**Stromal (n = 11)**		**Mixed (n = 15)**		**Regressive (n = 7)**	
	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile	Mean (Std)	25th / 75th Percentile
ADC (mono-exponential)	1.34×10^{-3} (4.09×10^{-4})	1.01×10^{-3}	1.02×10^{-3} (2.12×10^{-4})	0.80×10^{-3}	1.74×10^{-3} (2.4×10^{-4})	1.50×10^{-3}	1.42×10^{-3} (3.11×10^{-4})	1.11×10^{-3}	1.50×10^{-3} (4.18×10^{-4})	1.15×10^{-3}
D (IVIM)	1.24×10^{-3} (4.32×10^{-4})	0.91×10^{-3}	0.92×10^{-3} (1.93×10^{-4})	0.73×10^{-3}	1.65×10^{-3} (2.27×10^{-4})	1.40×10^{-3}	1.33×10^{-3} (2.98×10^{-4})	1.03×10^{-3}	1.42×10^{-3} (4.21×10^{-4})	1.08×10^{-3}
*D** (IVIM)	90.8×10^{-3} (10.9×10^{-2})	3.74×10^{-3}	69.2×10^{-3} ($7.72 \times 10^{-2)}$	6.71×10^{-3}	101×10^{-3} (14.3×10^{-2})	3.26×10^{-3}	58.3×10^{-3} (6.69×10^{-2})	3.89×10^{-3}	88.9×10^{-3} (4.90×10^{-2})	2.56×10^{-3}
f (IVIM)	0.18 (0.04)	0.10	0.15 (0.03)	0.10	0.18 (0.06)	0.10	0.16 (0.04)	0.09	0.18 (0.03)	0.07
f x *D** (IVIM)	1.41×10^{-2} (1.9×10^{-2})	5.67×10^{-4}	1.07×10^{-2} (1.18×10^{-2})	7.47×10^{-4}	2.18×10^{-2} (3.42×10^{-2})	4.63×10^{-4}	1.15×10^{-2} (1.92×10^{-2})	4.55×10^{-4}	1.32×10^{-2} (0.87×10^{-2})	3.71×10^{-4}

DDC (stretched exponential)	67.1 (1.77)	1.06×10^{-3}	8.79 (23.2)	0.79×10^{-3}	$1.15 \times 10^{+2}$ ($3.79 \times 10^{+2}$)	1.63×10^{-3}	12.9 (0.49)	1.16×10^{-3}	$4.68 \times 10^{+6}$ ($1.24 \times 10^{+7}$)	1.21×10^{-3}
α (stretched exponential)	0.83 (4.88×10^{-2})	0.75	0.76 (1.08×10^{-2})	0.70	0.88 (4.37×10^{-2})	0.83	0.85 (3.94×10^{-2})	0.79	0.87 (5.62×10^{-2})	0.82
D_k (kurtosis)	1.91×10^{-3} (3.79×10^{-4})	1.45×10^{-3}	1.89×10^{-3} (4.47×10^{-4})	1.21×10^{-3}	2.36×10^{-3} (4.15×10^{-4})	1.89×10^{-3}	2.03×10^{-3} (3.76×10^{-4})	1.48×10^{-3}	2.08×10^{-3} (4.82×10^{-4})	1.44×10^{-3}
K (kurtosis)	1.12 (0.45)	0.61	1.36 (0.32)	1.09	0.47 (0.13)	0.32	0.91 (0.34)	0.52	0.69 (0.30)	0.24

9.4.3: Wilms' Tumour Subtypes – One-way ANOVA Results

Five Wilms' tumour subtypes (blastemal, epithelial, stromal, mixed and regressive) were compared based on 9 diffusion parameters: ADC (mono-exponential), *D* (IVIM), *D** (IVIM), *f* (IVIM), *f* x *D** (IVIM), *DDC* (stretched exponential), α (stretched exponential), D_k (kurtosis), and *K* (kurtosis). For every viable and whole tumour ROI the 25th percentile of each parameter was calculated (apart from *f* where the 75th percentile was calculated). Mean and medians were also investigated however the 25th percentile provided more significant results.

9.4.3.1: One-way ANOVA Results - Mono-exponential

For the viable ROIs ADC had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance ($p = 0.05$). Significant differences between groups was found by a one-way ANOVA: $F(4, 42) = 9.42$, $p = 1.6 \times 10^{-5}$. Post-hoc Tukey tests revealed that there were significant differences between the blastemal and stromal ($p = 4.05 \times 10^{-4}$), epithelial and stromal ($p = 1.8 \times 10^{-5}$), mixed and stromal ($p = 0.003$) and regressive and stromal subtypes ($p = 0.024$), with the stromal being significantly higher than the other subtype groups.

For the whole tumour ROIs, ADC had no significant outliers, the data was normally distributed, but it did violate the assumption of homogeneity of variance ($p = 0.032$). For this reason a Welch ANOVA was used which showed significant differences between groups: $F(4, 42) = 7.24$, $p = 1.58 \times 10^{-4}$. Post-hoc Games-Howell tests revealed that there were similar significant differences to the viable ROI with the stromal group being significantly higher than the blastemal ($p = 0.02$), epithelial ($p = 4.0$ x 10-6), and mixed ($p = 0.002$). The difference between regressive and stromal, however, was not found to be significant ($p > 0.05$), unlike with the viable ROIs. However, epithelial ADC values were significantly smaller than the mixed which was not seen in the viable ROIs ($p = 0.021$). Figure 9.3 shows the ADC ANOVA results.

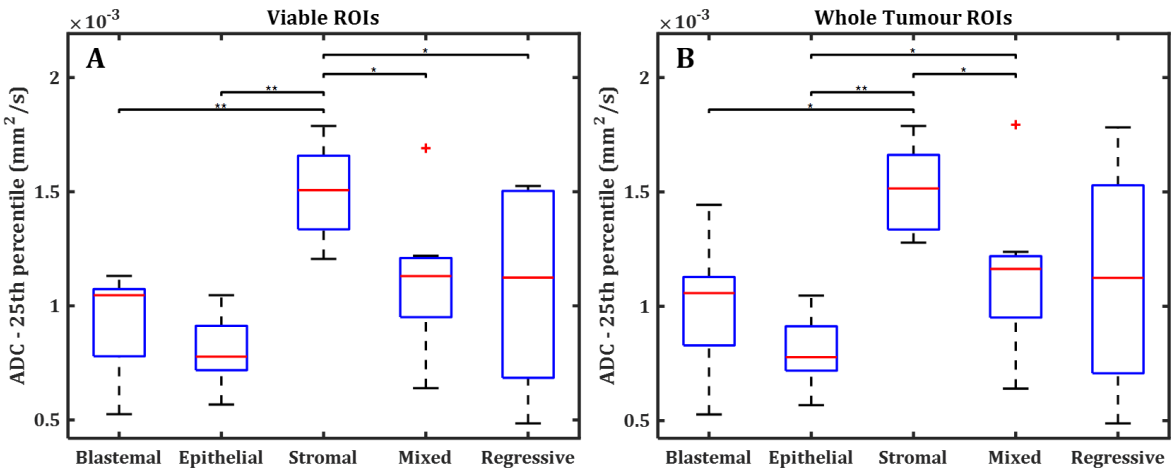

Figure 9.3. One-way ANOVA results for Wilms' tumour subtype separation using ADC 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

9.4.3.2: One-way ANOVA Results - IVIM

For the viable ROIs, *D* had no significant outliers, the data was normally distributed, and but it did violate the assumption of homogeneity of variance ($p = 0.042$). Significant differences between groups was found by a Welch ANOVA: $F(4, 42) = 10.36$, $p = 6.0 \times 10^{-6}$. Post-hoc Games-Howell tests revealed that there were significant differences between the blastemal and stromal ($p = 0.001$), epithelial and stromal ($p = 2.0 \times 10^{-6}$), mixed and stromal ($p = 0.001$) with stromal being significantly higher than the other subtype groups. Additionally, there was a significant difference between epithelial and mixed subtypes ($p = 0.016$), with mixed being significantly higher than epithelial subtypes.

For the whole tumour ROIs, *D* also had no significant outliers, the data was normally distributed, and violated the assumption of homogeneity of variance ($p = 0.014$). Significant differences between groups was found by a Welch ANOVA: $F(4, 42) = 7.89$, $p = 7.8 \times 10^{-5}$. Post-hoc Games-Howell tests revealed the same significant differences as *D* from the viable ROIs; with stromal being significantly higher than blastemal ($p = 0.014$), epithelial ($p = 2.0 \times 10^{-6}$), mixed ($p = 0.001$). The significantly higher *D* values were also seen in mixed compared to epithelial subtypes ($p = 0.013$). Figure 9.4 shows the *D* ANOVA results.

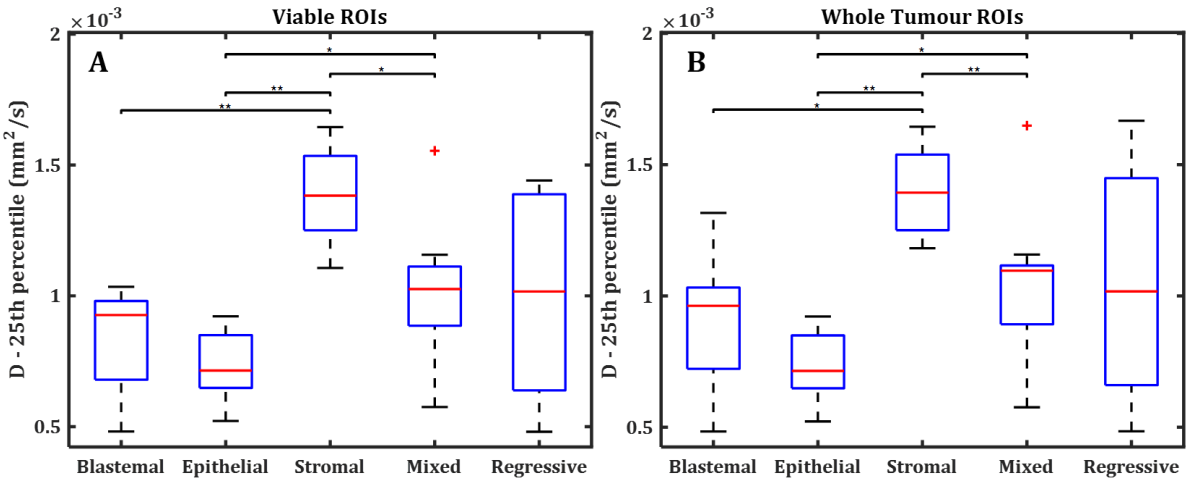

Figure 9.4. One-way ANOVA results for Wilms' tumour subtype separation using *D* 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

For the viable ROIs, *D** had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance ($p = 0.738$). A one-way ANOVA showed significant differences between groups: $F(4, 42) = 6.9$, $p = 2.32 \times 10^{-4}$. Post-hoc Tukey tests revealed that epithelial subtypes had significantly higher *D** values than blastemal ($p = 0.024$), stromal ($p = 0.001$), mixed ($p = 0.004$) and regressive ($p = 2.16 \times 10^{-4}$) subtypes.

For the whole tumour ROIs, *D** results were the same as with the viable ROIs. There were no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance ($p = 0.836$). One way ANOVA: $F(4, 42) = 7.65$, $p = 1.01 \times 10^{-4}$. Epithelial were shown, again, to have significantly higher *D** values than blastemal ($p = 0.007$), stromal ($p = 3.2 \times 10^{-4}$), mixed ($p = 0.002$) and regressive ($p = 8.6 \times 10^{-5}$) subtypes. Figure 9.5 shows the *D** ANOVA results.

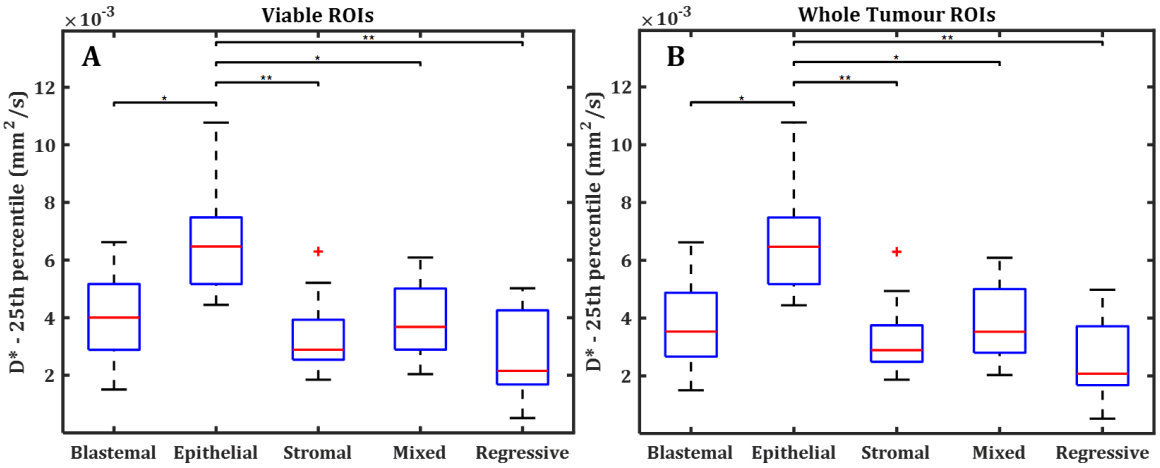

Figure 9.5. One-way ANOVA results for Wilms' tumour subtype separation using D^* 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

For *f*, both the viable and whole tumour ROIs revealed one significant outlier (an epithelial tumour). After removing this tumour from the cohort, both the viable and whole tumour ROIs had normally distributed data and did not violate the assumption of homogeneity of variance ($p = 0.16$: viable, $p = 0.11$: whole tumour). One-way ANOVAs showed no significant differences between groups for the viable ROIs ($p = 0.14$) or whole tumour ROIs ($p = 0.15$). Figure 9.6 shows the *f* ANOVA results.

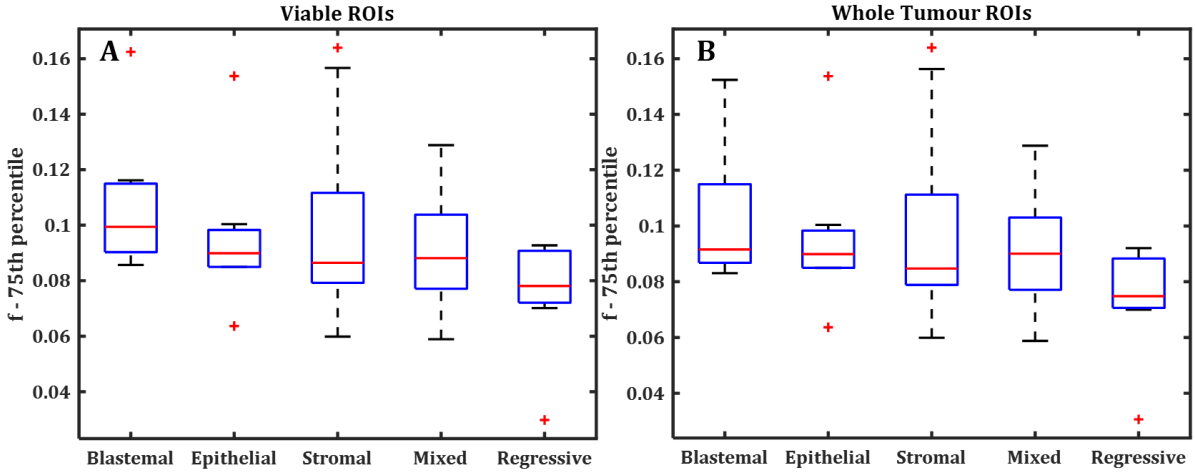

Figure 9.6. One-way ANOVA results for Wilms' tumour subtype separation using *f* 75th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

The *f* x D^* parameter was also investigated, and the viable ROIs had no significant outliers, the data was normally distributed, and it did not violate the assumption of

homogeneity of variance (p = 0.202). A one-way ANOVA revealed significant differences between groups: $F(4, 42) = 3.79$, p = 0.01. Post-hoc Tukey tests showed epithelial values to be significantly higher than stromal (p = 0.04), mixed (p = 0.021) and regressive (p = 0.016).

The same results were shown for the whole tumour ROI, with f x D^* having no significant outliers, normally distributed data, and not violating the assumption of homogeneity of variance (p = 0.20). A one-way ANOVA revealed significant differences between groups: $F(4, 42) = 4.12$, p = 0.007. Post-hoc Tukey tests, again, showed epithelial values to be significantly higher than stromal (p = 0.032), mixed (p = 0.016) and regressive (p = 0.007). Figure 9.7 shows the f x D^* ANOVA results.

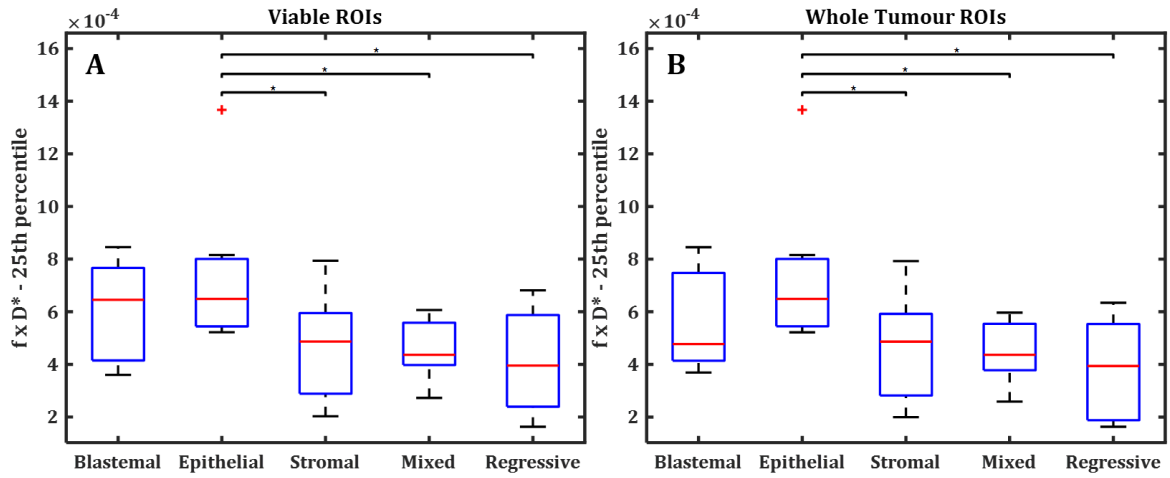

Figure 9.7. One-way ANOVA results for Wilms' tumour subtype separation using f x D^* 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

9.4.3.3: One-way ANOVA Results – Stretched Exponential

For the viable ROIs, *DDC* had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance (p = 0.116). A one-way ANOVA showed significant differences between groups: $F(4, 42) = 9.13$, $p = 2.1 \times 10^{-5}$. Post-hoc Tukey tests revealed that stromal subtypes had significantly higher *DDC* values compared to blastemal ($p = 4.6 \times 10^{-4}$), epithelial ($p = 2.6 \times 10^{-5}$), mixed (p = 0.004) and regressive subtypes (p = 0.027).

For the whole tumour ROIs, *DDC* also had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance (p = 0.076). A one-way ANOVA showed significant differences between groups: $F(4, 42) = 6.7$, $p = 2.89 \times 10^{-4}$. The same differences were found as with the viable ROIs with stromal values being significantly higher than blastemal (p = 0.014), epithelial ($p = 1.4 \times 10^{-4}$), mixed (p = 0.015). However, there was not a significant difference between stromal and regressive (p > 0.05) which was seen with the viable ROIs. Figure 9.8 shows the *DDC* ANOVA results.

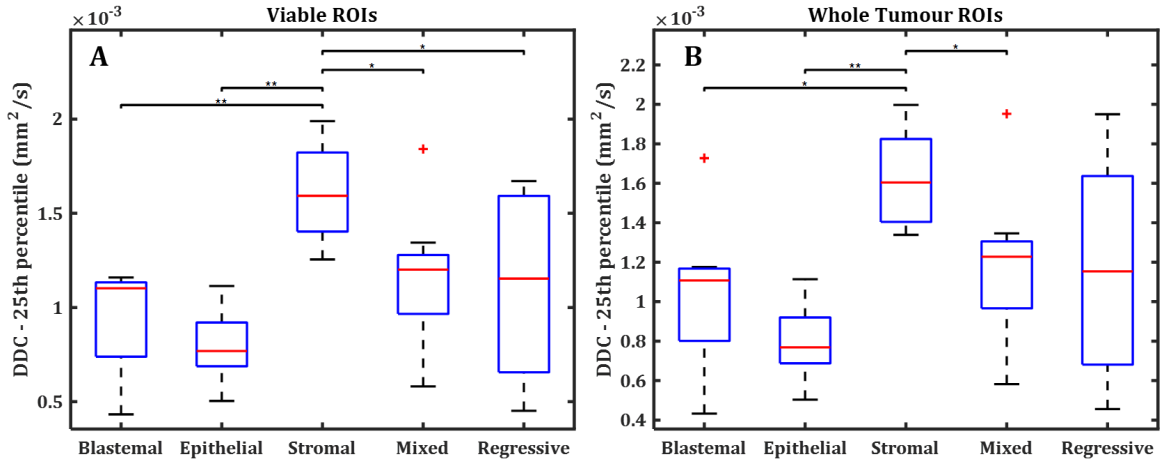

Figure 9.8. One-way ANOVA results for Wilms' tumour subtype separation using *DDC* 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

For the viable ROIs, α had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance (p = 0.901). A one-way ANOVA showed significant differences between groups: $F(4, 42) = 6.33$, $p = 4.41 \times 10^{-4}$. Post-hoc Tukey tests revealed that epithelial subtypes had significantly smaller α values compared to stromal (p = 0.001), mixed (p = 0.02) and regressive subtypes (p = 0.014). Additionally, stromal values were significantly greater than blastemal (p = 0.022).

For the whole tumour ROIs α also had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance (p = 0.807). A one-way ANOVA also showed significant differences between groups: $F(4, 42) = 6.05$, $p = 2.89 \times 10^{-4}$. Post-hoc Tukey tests showed that epithelial subtypes had significantly smaller α than stromal (p = 0.001), mixed (p = 0.017) and regressive subtypes (p = 0.008), which was also seen in the viable ROIs. However, no significant differences were found between stromal and blastemal subtypes (p > 0.05), unlike the viable ROIs. Figure 9.9 shows the α ANOVA results.

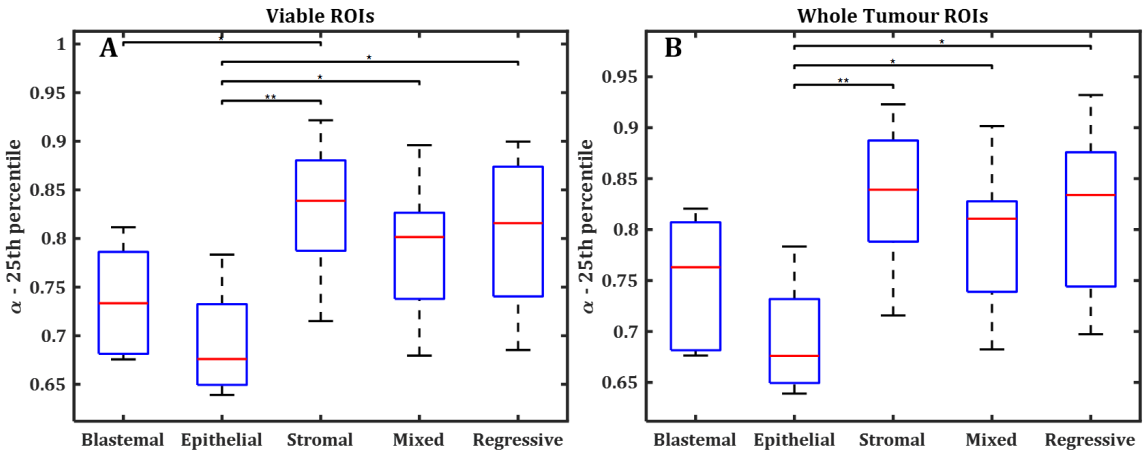

Figure 9.9. One-way ANOVA results for Wilms' tumour subtype separation using α 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

9.4.3.4: One-way ANOVA Results – Kurtosis

For $D_{k,}$ both the viable and whole tumour ROIs revealed one significant outlier each: two different blastemal tumours (viable – *25th percentile:* 7.1 x 10^{-4}, whole tumour – *25th percentile*: 2.31 x 10^{-3}). Once these tumours were removed from the cohort, both the viable and whole tumour ROIs had normally distributed data and did not violate the assumption of homogeneity of variance (p = 0.051: viable, p = 0.131: whole tumour). For the viable ROI a one-way ANOVA revealed significant differences between groups F (4, 41) = 5.25, p = 0.002. Post-hoc Tukey tests showed that stromal subtypes had higher $D_{k,}$ values than epithelial (p = 0.002), mixed (p = 0.020), and regressive (p = 0.025) subtypes. The whole tumour ROIs also revealed significant group differences F (4, 41) = 4.71, p = 0.003. However, post-hoc Tukey results showed a different pattern; there were no significant differences between stromal and epithelial, regressive or mixed subtype ($p > 0.05$). However, there were significant differences between blastemal and regressive (p = 0.024), with blastemal having higher values, and epithelial and regressive (p = 0.004), with regressive having higher values. Figure 9.10 shows the D_k ANOVA results.

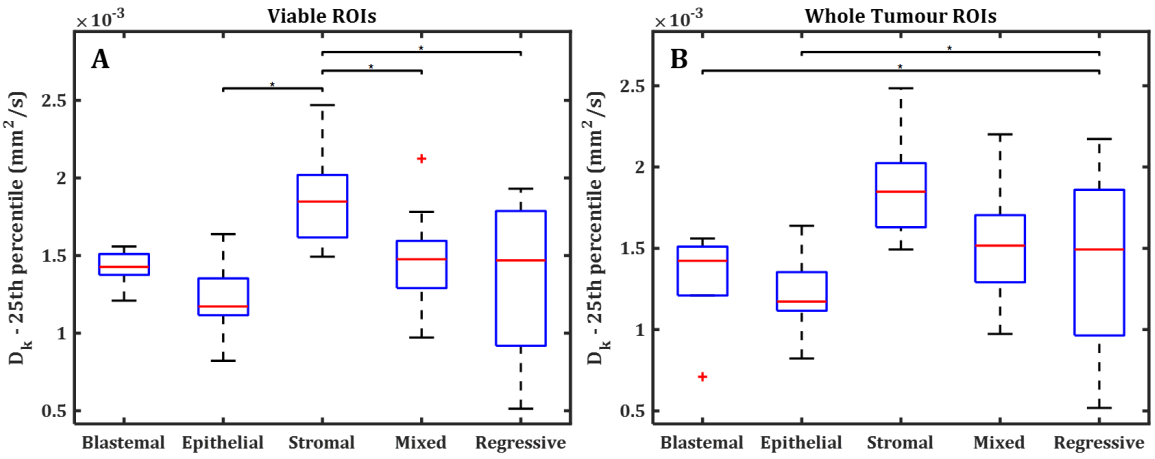

Figure 9.10. One-way ANOVA results for Wilms' tumour subtype separation using D_k 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

For the viable ROIs, *K* had no significant outliers, the data was normally distributed, and it did not violate the assumption of homogeneity of variance ($p = 0.657$). A one-way ANOVA revealed significant differences between groups $F(4, 42) = 12.07$, $p = 1.0 \times 10^{-6}$. Post-hoc Tukey tests showed that epithelial values were significantly higher than blastemal ($p = 0.036$), stromal ($p = 3.0 \times 10^{-6}$), regressive ($p = 6.0 \times 10^{-6}$), and mixed ($p = 2.93 \times 10^{-4}$). In addition, blastemal values were also shown to be higher than regressive, and this difference just reached significance ($p = 0.047$). Blastemal values were also higher than stromal values, but this just failed to reach significance ($p = 0.053$).

For the whole tumour ROIs, *K* had one significant outlier: a mixed tumour (*25th percentile:* 1.23). After removing this tumour from analysis, the data was normally distributed, and it did not violate the assumption of homogeneity of variance ($p = 0.196$). A one-way ANOVA revealed significant differences between groups $F(4, 41) = 13.21$, $p = 5.32 \times 10^{-7}$. Post-hoc Tukey tests showed similar results to the viable ROIs; epithelial values were significantly higher than blastemal ($p = 0.009$), stromal ($p = 1.0 \times 10^{-6}$), regressive ($p = 2.0 \times 10^{-6}$), and mixed ($p = 3.4 \times 10^{-5}$). The difference between blastemal and regressive was, however, not significant ($p > 0.05$), unlike with the viable ROIs. Figure 9.11 shows the *K* ANOVA results.

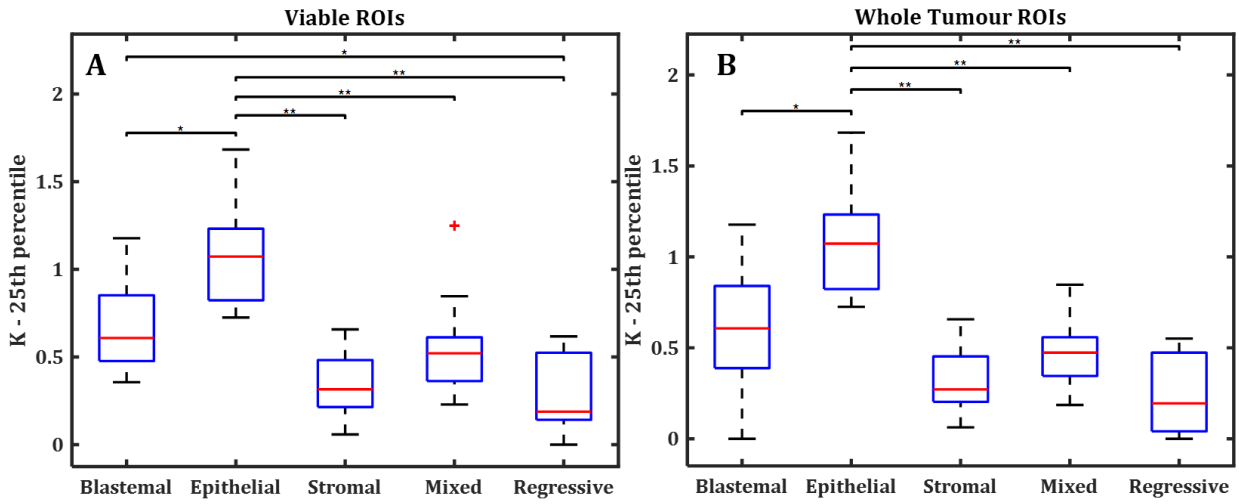

Figure 9.11. One-way ANOVA results for Wilms' tumour subtype separation using K 25th percentile as the dependent variable. **A:** Data acquired with viable ROIs. **B:** Data acquired with whole tumour ROIs. Post-hoc significant differences are shown by bars, * = $p < 0.05$, ** = $p \leq 0.001$

9.5: Discussion

This study sought to determine whether various diffusion parameters could distinguish Wilms' tumour subtypes from one another. Nine post-chemotherapy diffusion parameters were investigated: ADC (mono-exponential), D (IVIM), D^* (IVIM), f (IVIM), f x D^* (IVIM), *DDC* (stretched exponential), α (stretched exponential), D_k (kurtosis), and K (kurtosis). Necrotic portions of the tumour were excluded according to results from Chapter 8. Additionally, whole tumour volume ROIs were also investigated to determine whether viable ROIs improved subtype identification. It was found that there were significant differences between subtypes when compared using the 25th percentile of each of the parameters for both whole volume ROIs and viable ROIs, except for f which did not reveal a significant result. Overall, slightly more differences were seen when using viable ROIs compared to whole tumour volume ROIs; for *DDC* a regressive-stromal difference was seen, for α a blastemal-stromal difference was seen, and for K a blastemal-regressive difference was revealed, whereas these were not present when using whole tumour volume ROIs. As necrosis identification is also a very important aspect of Wilms' tumour treatment response and outcome, it is suggested that one should use the enhancement threshold derived in Chapter 8 (> 0.33 = viable) when making subtype separations.

Previous work showed ADC values of the blastemal subtype (n = 5) to be significantly smaller than the ADC values of mixed (n = 11), stromal (n = 6) and regressive

(n = 5) subtypes, and also epithelial (n = 4) ADC values to be smaller than stromal values[3]. In the present study, while blastemal ADC values were lower than mixed and regressive this difference was not significant, however, both blastemal and epithelial ADC values were significantly smaller than stromal. The reason for these differences may be due to the increased number of each subtype in the present study, and therefore a greater spread of ADC values leading to less significant differences between the blastemal and other subtypes. However, as the blastemal ADC values were still low in the present study, it supports previous ideas that blastemal tissue which remains after treatment is chemotherapy-resistant, with a high cellularity[9]. There were more significant differences seen between stromal and other subtypes compared to the previous study[3], with all other subtypes having significantly smaller ADC values when viable ROIs were used. It has been suggested that chemotherapy might induce maturation of Wilms' tumour tissue resulting in stromal subtypes[9], therefore higher ADC values may reflect the decrease in cellularity from this maturation process. Additionally, previous research has shown that median post-chemotherapy ADC values correlated with the proportion of stromal tissue and that these ADC values were significantly higher than other subtypes[4], further details on this investigation can be found in Chapter 4.2.4.

For the IVIM model, no significant differences between subtypes were found when using the *f* parameter based on the 75th percentile. This metric was chosen as lower values of *f* are linked to necrosis and this study aimed to exclude necrotic regions as to not obscure values. It has been previously shown that *f* was significantly lower in necrotic liver tumours compared to the viable tumour components as defined via gadolinium enhancement and histological assessment[8]. Despite excluding necrotic tissue in the present study using a threshold and also taking higher values to further avoid necrotic regions, no subtype differences were seen.

*D** revealed many significant differences; with the epithelial subtype having higher values than all other subtypes. This is a very interesting result as previously epithelial values have appeared very similar to blastemal, with both showing higher cellularity and lower ADC values[3]. Previous work using IVIM in the kidney has shown *D** to be of less value compared to *D* and *f.* For example, no differences were seen when comparing diabetic kidney tissue to controls in *D** unlike *D* and *f*[10], and also *D* and *f* have been shown to be related to estimated glomerular filtration rate (eGFR) unlike *D**[11]. However, there are a few studies which have shown *D** to provide useful information with it being able to detect significant differences between normal kidney function and mild impairment, and between normal and severe impairment in the renal cortex, whereas *D* and *f* could not

make these distinctions[12]. *D** is thought to be related to perfusion in the randomly orientated capillaries, however other factors may influence this parameter such as tubular flow or glandular secretion[13]. Epithelial tumours are made up of tubular structures[14], therefore the higher *D** values seen in the epithelial subtypes may reflect the flow through these tubules, however this is just speculation. Previously, *D** has been shown to have low reproducibility[15–20] (Chapter 6), and high variability, particularly in hypo-perfused tissue[21]. Therefore, although *D** could be very useful at separating epithelial and blastemal subtypes, this distinction may not appear if different acquisitions are used, thus, limiting the clinical transference of this finding. This result therefore suggests that non-Gaussian models may provide additional clinical information compared to ADC, but optimisation of the models to ensure low variability is needed before making definitive conclusions.

Out of the four models, the stretched exponential model was previously shown to provide the best fit to the Wilms' tumour data (Chapter 7), therefore it could have been expected to provide more subtype distinctions compared to ADC. However, *DDC* revealed the same results as ADC for subtype separations. On the other hand, α revealed slightly different results; with less stromal differences (only epithelial-stromal and blastemal-stromal were seen). However, more epithelial differences were observed, unlike ADC, with epithelial having significantly smaller α values compared to stromal, mixed and regressive. Therefore, although the model may provide a better fit to the data (Chapter 7) and α has been shown to be reproducible[17,18,22] (Chapter 6), one needs to determine which subtype differences are the most important; as ADC can provide many differences from the stromal subtype and α from the epithelial subtype. In terms of treatment and outcome, in most cases, only blastemal would differ from the other subtypes[1]. Therefore, it may not be necessary to acquire multiple b values to obtain α.

For the kurtosis model, different outliers were identified in D_K depending on which ROIs were used. In addition, different subtype distinctions were also revealed depending on ROIs. Whole volume ROIs could distinguish blastemal and epithelial from regressive, and viable ROIs could separate stromal from epithelial, mixed and regressive. The viable ROIs provided a similar pattern to the other diffusion parameters (*DDC*, *D*, *ADC*) and are more likely to reflect the differences in the viable tumour components compared to whole tumour ROIs. Furthermore, more subtype separations were identified using the viable ROIs with *K* compared to the whole tumour ROIs. *K* provided many subtype separations and could distinguish epithelial from blastemal which as previously mentioned has not been shown before[3]. *K* provided the same results to *D**, with the addition of blastemal also showing higher values than regressive. It has been suggested that a higher value of *K* may

imply a more complex tissue environment[23]. It could be that the epithelial subtypes have a more complex environment as they are highly cellular at the surface of the tubular structures and then are less cellularly dense in the centre of these structures, whereas the blastemal subtypes are more uniform; consisting of densely packed cells[14]. *K* may also reflect the interaction of water molecules with intracellular compounds and cell memebranes[7,24], however, it is unclear what exactly *K* describes in a biophysical sense. Despite the potentially useful subtype separations (with blastemal distinctions), *K* has been shown have high variability[17] and low reproducibility (Chapter 6). This may be due to the fitting methods where without upper bounds *K* produces very high values, and when boundaries are applied *K* simply hits this limit (this is discussed further in Chapter 6.5). Therefore, as with *D*,* it is important to be aware of the variability of *K,* however this parameter did show promising results.

There are several limitations to this study, firstly it relies on the threshold generated in Chapter 8 to accurately identify necrotic tissue. Additionally, Wilms' tumours are very heterogeneous, and it might be considered crude to label a tumour as a certain subtype when various regions of different cellular compositions exist within it. Future analysis should consider reviewing each MRI slice and matching it to individual histological slices to gain a better understanding of how diffusion parameters describe the tissue. Previous work has shown this is feasible and that the 25th percentile of ADC was correlated with the percent of blastemal tissue[25]. However, for the present study this was not possible and so the SIOP-2001 categorisation system was used.

The 25th percentile was used as the dependent variable to avoid areas with high diffusion values. Previous work has found this to be a useful metric in subtype distinctions[3,4,25], however it is possible that other metrics may also provide useful subtype separation.

Finally, diffusion measurements were generated from 7 and 8 b value ranges. These were combined to ensure a larger cohort. As shown in a previous chapter (Chapter 6) the coefficients of variation (CV) of *DDC* (CV: 116.4%), *K* (CV: 31.82%), and *D** (CV:27.87%) were very high when these b value ranges were compared in healthy adult renal tissue. Therefore, one needs to be cautious when interpreting the results as the same differences may not be found when different b values are used.

9.6: Conclusions and Summary

Wilms' tumour subtypes are identified via histological analysis after surgery, therefore this study investigated whether diffusion parameters could non-invasively distinguish Wilms' tumour subtypes prior to surgery, as this may impact on treatment decisions.

Nine diffusion parameters were investigated, and it was expected that all of these would provide a significant one-way ANOVA result. However, the *f* parameter did not reach significance unlike the other eight parameters which found significant differences when using the 25th percentile.

As hypothesised, ADC values were significantly higher in stromal tissue compared to other subtypes ($p < 0.05$). ADC values of blastemal and epithelial subtypes were also lower than the other subtypes as predicted, but this relationship did not reach significance.

Most distinctions were seen when using non-Gaussian models as hypothesised. *D** (IVIM) and *K* (kurtosis) provided the most separations and managed to distinguish between blastemal and epithelial subtypes, which previously has not been possible. However, both of these parameters have low reproducibility and therefore these results may not be replicated in future studies. Therefore, this result should be interpreted as a suggestion that non-Gaussian models may be clinically useful in subtype identification, but more work is needed to optimise the reproducibility of certain parameters.

As predicted, excluding necrotic regions prior to separating subtypes improved identification in some parameters. As necrosis identification is important in assessing treatment and predicting patient outcomes, it is suggested that this tissue is identified and removed for subtype identification.

Main finding: For subtype identification in Wilms' tumour, analysis should be limited to the viable portions of the tumour. There is potential for non-Gaussian models (IVIM and kurtosis) to be able to identify subtypes, and distinguish between blastemal and epithelial subtypes, however greater investigation is needed to optimise and validate these parameters in terms of reproducibility.

References: Chapter 9

1. Vujanić, G. M. *et al.* Revised International Society of Paediatric Oncology (SIOP) working classification of renal tumors of childhood. *Med. Pediatr. Oncol.* **38**, 79–82 (2002).
2. Humphries, P. D., Sebire, N. J., Siegel, M. J. & Olsen, Ø. E. Tumors in Pediatric Patients at Diffusion-weighted MR Imaging: Apparent Diffusion Coefficient and Tumor Cellularity. *Radiology* **245**, 848–854 (2007).
3. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
4. Littooij, A. S., Sebire, N. J. & Olsen, Ø. E. Whole-tumor apparent diffusion coefficient measurements in nephroblastoma: Can it identify blastemal predominance? *J. Magn. Reson. Imaging* **45**, 1316–1324 (2017).
5. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).
6. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
7. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
8. Wagner, M. *et al.* Diffusion-weighted MR Imaging for the Regional Characterization of Liver Tumors. *Radiology* **264**, 464–472 (2012).
9. Weirich, A. *et al.* Clinical impact of histologic subtypes in localized non-anaplastic nephroblastoma treated according to the trial and study SIOP-9/GPOH. *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol.* **12**, 311–319 (2001).
10. Deng, Y. *et al.* Use of intravoxel incoherent motion diffusion-weighted imaging to detect early changes in diabetic kidneys. *Abdom. Radiol.* 1–6 (2018). doi:10.1007/s00261-018-1521-4
11. Mao, W. *et al.* Chronic kidney disease: Pathological and functional evaluation with intravoxel incoherent motion diffusion-weighted imaging. *J. Magn. Reson. Imaging JMRI* **47**, 1251–1259 (2018).

12. Ichikawa, S. *et al.* Intravoxel incoherent motion imaging of the kidney: alterations in diffusion and perfusion in patients with renal dysfunction. *Magn. Reson. Imaging* **31**, 414–417 (2013).
13. Koh, D.-M., Collins, D. J. & Orton, M. R. Intravoxel Incoherent Motion in Body Diffusion-Weighted MRI: Reality and Challenges. *Am. J. Roentgenol.* **196**, 1351–1361 (2011).
14. Histopathology images of Wilms tumor (Nephroblastoma) by PathPedia.com: Pathology e-Atlas. Available at: https://www.pathpedia.com/education/eatlas/histopathology/kidney/wilms_tumor_(nephroblastoma).aspx. (Accessed: 17th December 2018)
15. Barbieri, S., Donati, O. F., Froehlich, J. M. & Thoeny, H. C. Comparison of Intravoxel Incoherent Motion Parameters across MR Imagers and Field Strengths: Evaluation in Upper Abdomina Organs. *Radiology* **279**, 784–794 (2015).
16. Cui, Y., Dyvorne, H., Besa, C., Cooper, N. & Taouli, B. IVIM diffusion-weighted imaging of the liver at 3.0 T: Comparison with 1.5 T. *Eur. J. Radiol. Open* **2**, 123–128 (2015).
17. Jerome, N. P. *et al.* Repeatability of derived parameters from histograms following non-Gaussian diffusion modelling of diffusion-weighted imaging in a paediatric oncological cohort. *Eur. Radiol.* **27**, 345–353 (2017).
18. Reischauer, C., Patzwahl, R., Koh, D.-M., Froehlich, J. M. & Gutzeit, A. Non-Mono-Exponential Analysis of Diffusion-Weighted Imaging for Treatment Monitoring in Prostate Cancer Bone Metastases. *Sci. Rep.* **7**, (2017).
19. Sun, H., Xu, Y., Xu, Q., Shi, K. & Wang, W. Rectal cancer: Short-term reproducibility of intravoxel incoherent motion parameters in 3.0T magnetic resonance imaging. *Medicine (Baltimore)* **96**, e6866 (2017).
20. Jakab, A., Tuura, R., Kottke, R., Kellenberger, C. J. & Scheer, I. Intra-voxel incoherent motion MRI of the living human foetus: technique and test–retest repeatability. *Eur. Radiol. Exp.* **1**, (2017).
21. Meeus, E. M. *et al.* Evaluation of intravoxel incoherent motion fitting methods in low-perfused tissue. *J. Magn. Reson. Imaging* **45**, 1325–1334 (2017).
22. Winfield, J. M. *et al.* Modelling DW-MRI data from primary and metastatic ovarian tumours. *Eur. Radiol.* **25**, 2033–2040 (2015).
23. Steven, A. J., Zhuo, J. & Melhem, E. R. Diffusion Kurtosis Imaging: An Emerging Technique for Evaluating the Microstructural Environment of the Brain. *Am. J. Roentgenol.* **202**, W26–W33 (2013).

24. Le Bihan, D. Apparent diffusion coefficient and beyond: what diffusion MR imaging can tell us about tissue structure. *Radiology* **268**, 318–322 (2013).
25. Littooij, A. S. *et al.* Apparent diffusion coefficient as it relates to histopathology findings in post-chemotherapy nephroblastoma: a feasibility study. *Pediatr. Radiol.* **47**, 1608–1614 (2017).

Chapter 10: Thesis Conclusions and Summary

The primary aim of this thesis was to investigate the use of diffusion weighted imaging in Wilms' tumours. Currently in the UK, in most cases, Wilms' tumour is treated with pre-operative chemotherapy and then surgery to remove the tumour, or the entire kidney[1]. Most patients will have MRI scans prior to chemotherapy and post-treatment, prior to surgery. DWI is often obtained as part of the standard protocol for Wilms' tumour, and research has investigated the use of ADC in terms of separation from benign tumours[2,3], separation from neuroblastoma[4], chemotherapy changes[5–7] and subtype identification[7–9]. This thesis aimed to build on previous research by using non-Gaussian models of diffusion: IVIM[10], stretched exponential[11] and kurtosis[12]. Although multiple b values are required for these models, this thesis explored whether the models would better describe the diffusion signal and provide additional information regarding tissue status, thus justifying the longer scan time.

This thesis consisted of four experimental chapters (chapters 6 – 9) which investigated the reproducibility of diffusion parameters in healthy adult kidney tissue, the fitting of the diffusion models in Wilms' tumour, the ability to identify necrotic tissue without using gadolinium, and subtype identification.

10.1: Chapter Summaries

10.1.1: Summaries: Chapter 6 Summary

The first study investigated the reproducibility of eight diffusion parameters from four models of diffusion in five pairs of healthy adult kidneys. Coefficients of variation and Bland-Altman analyses investigated the variability in parameters across magnetic field strengths (1.5T vs. 3T) and b values. Low variability was defined as coefficients of variation and levels of bias (as assessed by the mean difference from Bland-Altman analysis): ≤15%.

Aims and Hypotheses:

- In healthy adult kidneys there will be a difference in reproducibility between the derived parameters from each model.

- For both experiments (field strength and b values) it is expected that the greatest variability will be seen in D^* and f (IVIM) and K (kurtosis).
- The other parameters ADC (mono-exponential), D (IVIM), DDC and α (stretched exponential) and D_k (kurtosis) will have lower variability.

Conclusions

- Differences in reproducibility were found between the parameters.
- The greatest variability was found in D^* and K as expected, however f was found to be reproducible. DDC, was also found to be highly variable.
- ADC, D, D_k, and α were all found to have low variability

Main Finding: Most parameters (ADC, D, D_k, and α) are reproducible across field strengths and when different b values are acquired. More caution needs to be taken when interpreting D^*, K and DDC as these parameters have high levels of variability.

10.1.2: Summaries: Chapter 7 Summary

The second study tested the how well each of the four diffusion models described the raw diffusion data in Wilms' tumours, and the contralateral unaffected normal kidney. Goodness of fit was based on the Akaike information criterion.

Aims and Hypotheses:

- In contralateral renal tissue IVIM will provide a superior fit to the raw diffusion data compared to the mono-exponential model.
- As there is a lack of research into stretched exponential and kurtosis models in normal renal tissue and a lack of research into all three non-Gaussian models in Wilms' tumour, specific directional hypotheses cannot be made on these models. This study thus aims to identify which of the models provide the best fit to this data.
- It is expected that the different Wilms' tumour subtypes will favour different models as they have different cellular environments.

Conclusions:

- In the contralateral renal tissue, the IVIM model provided the best fit to the diffusion data
- The stretched exponential model provided the best fit to the Wilms' tumour data

- While majority of tumours favoured the stretched exponential model, there was high heterogeneity within the tumours, and kurtosis and IVIM models also provided good fits to the data
- There were no differences in model preferences based on subtype

Main Finding: Non-Gaussian models provide superior fits to diffusion data in Wilms' tumour and contralateral normal kidney tissue compared to a mono-exponential model. Out of 112 tumours only 1 favoured the mono-exponential model, and out of 77 contralateral kidneys none favoured the mono-exponential model.

10.1.3: Summaries: Chapter 8 Summary

The third study attempted to identify necrotic tissue within post-chemotherapy Wilms' tumours using DWI and T_1w imaging. The volume of necrosis in a Wilms' tumour is informative of treatment response and survival rates[13], and currently can be visualised as non-enhancing tissue using T_1w after injecting a gadolinium-based contrast agent. However, there are many associated risks with gadolinium and it may not always be a viable option for all patients[14].

Aims and Hypotheses:

- There will be a relationship between mean diffusion values and mean fractional enhancement (based on gadolinium contrast) in Wilms' tumours. An inverse relationship will be seen in the majority of parameters with an increase in ADC (mono-exponential), *D* and *D**(IVIM), *DDC* and α (stretched exponential), D_k and *K* (kurtosis) being significantly ($p < 0.05$) related to a decrease in mean fractional enhancement, the parameter f x D^* (IVIM) will also be investigated and the same relationship will be seen.
- A positive relationship will be seen with *f*, with an increase in mean *f* being significantly ($p < 0.05$) related to an increase in mean fractional enhancement.
- The IVIM parameters are expected to provide the highest adjusted R^2 values.
- The selected model from the regression analysis will be compared to the gadolinium method and will have little bias. A small bias will be defined as below 10% based on Bland-Altman analysis.

- A threshold will be determined which can separate necrotic and viable Wilms' tumour tissue based on diffusion and T_1weighted imaging

Conclusions:

- Multiple linear regressions revealed that combining a single diffusion parameter and T_1w (without contrast) was significantly related to fractional enhancement in Wilms' tumours.
- There was a significant inverse relationship between mean ADC (mono-exponential), *D*, *D**, and *f* x *D* (IVIM), *DDC* and α (stretched exponential), D_k and *K* (kurtosis) and mean fractional enhancement. There was also a significant positive relationship between mean *f* and mean fractional enhancement
- The *D*-T_1w_{norm} model (*D* from IVIM) provided the highest adjusted R^2 value. However, this was not much greater than ADC, and as ADC is more widely available the ADC-T_1w_{norm} model was selected.
- The ADC-T_1w_{norm} model provided visually similar maps to the gadolinium method and a small bias was found between the two methods (9%)
- A threshold off 33% was generated, thus meaning that tissue with signal enhancement of less than 33% (using conventional, post-gadolinium T_1w data, or using the ADC-T_1w_{norm} model) could identify necrotic tissue with high specificity (90%) and sensitivity (85%)

Main Finding: ADC and T_1w imaging can provide an alternative approach to identify and quantify necrosis in Wilms' tumour without the need for contrast agents

10.1.4: Summaries: Chapter 9 Summary

The final study aimed to identify Wilms' tumour subtypes using DWI. Currently subtypes can only be identified via histological analysis post-surgery, therefore this study attempted to non-invasively identify the subtypes using DWI. Necrotic tissue was removed using the threshold developed in the third study (Chapter 8) to avoid obscuring the distribution of diffusion measurements within the tumour volume. One-way ANOVAs compared five subtypes: epithelial, stromal, mixed, and regressive (all intermediate risk) and blastemal (high risk), using the 25th percentile of diffusion parameters from the four diffusion models (and the 75th percentile for the *f* parameter from IVIM).

Aims and Hypotheses:

- There will be significant differences between the subtypes based on diffusion values.
- The non-Gaussian models will provide more subtype separations compared to the mono-exponential model
- The blastemal and epithelial subtypes will have lower ADC values compared to the other subtypes.
- The stromal subtype will have the highest ADC values compared to the other subtypes.
- Limiting the analysis to just the viable tumour tissue will provide more significant differences between the subtypes compared to using the whole tumour values.

Conclusions:

- Out of nine parameters investigated all but one (*f*) revealed significant ANOVAs between subtypes.
- *D** and *K* from the non-Gaussian models (IVIM and kurtosis) provided the best subtype stratification, and could separate epithelial and blastemal subtypes, which has not been shown previously[7].
- The blastemal and epithelial subtypes had lower ADC values although this did not reach significance when compared to all other subtypes.
- The stromal subtype had significantly higher ADC values compared to the other subtypes.
- Limiting the analysis to the viable portions of the tumour provided more subtype distinctions compared to whole tumour analysis.

Main finding: For subtype identification in Wilms' tumour, analysis should be limited to the viable portions of the tumour. There is potential for non-Gaussian models (IVIM and kurtosis) to be able to identify subtypes, and distinguish between blastemal and epithelial subtypes, however greater investigation is needed to optimise and validate these parameters in terms of reproducibility.

10.1.5: Overall Summary

Overall this thesis has explored the feasibility and utility of acquiring more advanced measures of diffusion (beyond ADC) in Wilms' tumour. It has shown that using

non-Gaussian models can provide better descriptions of diffusion data than a mono-exponential model, and that some of these parameters could be useful in clinical investigations such as identifying Wilms' tumour subtypes. It has also demonstrated that necrotic tissue can be identified using standard ADC and T_1w imaging with comparable results to gadolinium contrast-enhanced imaging.

10.2: Limitations

Limitations of each experiment are given in the discussion sections of each experimental chapter. Here, the overall limitations relating to the thesis as a whole are discussed.

The main limitation with this thesis, and clinical investigations in general, is the lack of control over the data. As the Wilms' DWI data was retrospectively collected and acquired using the standard clinical protocols at this institution, the MRI sequence parameters could not be adapted. This meant there was variation in b values used, and these were not optimised for the non-Gaussian models used throughout this thesis, further details on this are discussed in section 10.5.

Another limitation with using a clinical population is that it is very difficult to acquire a perfect sample. One cannot control the size and subtype of the tumours, or whether there is equal pre- and post-chemotherapy data, or whether all patients have histological analysis, or all the necessary MRI sequences. This can create challenges, for example when comparing subtypes with small uneven populations in each group.

A limitation which has been frequently refereed to throughout this thesis is the lack of a gold standard, and/or histological reference. As mentioned previously, currently only a subsection of the tumour is histologically sampled and a crude estimate of necrosis and main cellular subtype is given. As Wilms' tumours are very heterogenous this poses many challenges. Ideally the tumour would be entirely sampled, and each slice would be analysed, thus providing information about distinct populations of cellular environments throughout the tumour. Following this process, MRI slices could be matched to histology. This would improve MRI analysis as one would then be more confident in defining a DWI threshold for subtypes or necrosis identification.

10.3: Improvements to Experiments Conducted

From reviewing this thesis there are areas which could have been improved upon. Firstly, in Chapter 6 more volunteers could have been recruited. This would have increased the generalisability of the findings, however due to time constraints this was not feasible. Furthermore, for field strength comparisons sequences could have been adapted so that they were perfectly matched, with the only altered variable being the field strength. This would have enabled more definitive conclusions to have been drawn regarding reproducibility relating to field strength. The initial aim was to focus on the clinical sequences used at this institution as Wilms' data was to be retrospectively collected and therefore analysis would have to rely on the sequences which are commonly acquired here. Therefore, the sequences were not adapted.

A larger cohort of Wilms' patients would have improved the investigations in this thesis. A larger variety of patients with varying levels of treatment response would have assisted with necrosis identification, and greater numbers of epithelial and blastemal subtypes would have improved the subtype identification. Furthermore, it would have been beneficial to have all the MRI scans for every patient, and a detailed histology and notes regarding their diagnosis, and treatment. This would have helped build up a clearer picture of each patient, however due to data access this was not possible.

10.4: Implications for Future Research and Patient Care

This thesis has shown the potential of using ADC and non-Gaussian models of diffusion in Wilms' tumour. It has demonstrated that non-Gaussian models describe the raw diffusion signal better than a mono-exponential model, according to AIC values. This finding will help with interpretation of past and future research which relies on ADC, as alternative models may provide a more accurate representation of the underlying signal.

A number of the parameters from these non-Gaussian models were also shown to have low variability in normal renal tissue and given that they fit the data well, this may help with increasing the use of these models. Different centres will be able to combine their data and be confident that the parameters will be not vary based on sequence alterations. Additionally, certain parameters; D^*, K, and DDC, were not found to be reproducible and thus this will also assist with interpreting research and deciding whether results could be replicated.

Identifying subtypes *in vivo* would greatly improve patient care. Different subtypes will have different treatment paths, and thus early identification of these subtypes would improve personalisation of treatment. It would assist with decisions in increasing treatment in those with high risk subtypes and also avoiding unnecessary treatment in those with lower risk subtypes. Given the side effects and later effects of chemotherapy, defining a treatment path early would be greatly beneficial for this paediatric population.

This thesis demonstrated that ADC and parameters from non-Gaussian diffusion models have the potential to non-invasively identify subtypes. In line with previous research[7–9] this thesis has shown that the stromal subtype has a lower cellularity, represented by a high ADC, and that blastemal and epithelial subtypes have a higher cellular density indicated by a lower ADC. Previously, however blastemal and epithelial subtypes have not been able to be distinguished using ADC due to their similar cellular densities[7]. Epithelial subtypes are classed as intermediate risk and blastemal as high risk; thus, they would confer a different treatment path. Non-Gaussian models produce a range of parameters that may assist with this separation. This thesis suggested that by using the 25^{th} percentile of D^* and K there is a significant different between these two subtypes. However, given the variability of these parameters and the low number of tumours in each subtype this result needs to be further explored, which is discussed in section 10.5. However, this result does indicate the potential of these non-Gaussian models and how they may have clinical significance and thus paves the way for future research in this area.

A main clinical implication from this thesis was the ability to identify necrotic tissue without the need of a contrast agent. As discussed throughout this thesis there are associated challenges and risks with using contrast agents and thus an alternative approach would be beneficial. The proposed method uses current clinical sequences, thus easing clinical transference. Furthermore, it provided a threshold which would allow quantification of the volume of necrotic tissue, and this could be identified using either post-gadolinium T_1w data or predicted enhancement maps (using ADC and pre-non-contrast enhanced T_1w imaging). The volume of necrosis is related to treatment response[15] and patient outcome[13] and thus this threshold could be very clinically useful. Additionally, the utility of non-Gaussian models for detecting necrotic tissue was also investigated. However, they did not greatly improve this estimation, therefore although they may be superior at describing the Wilms' tumour diffusion data, they are not necessary for this particular clinical investigation. This may help with using this method in future research as ADC is a very common parameter and does not require multiple b values.

10.5: Future Work

Based on the findings of this thesis future work should focus on the optimisation of non-Gaussian models, in the hope of decreasing variability and increasing reproducibility. This could be done by exploring different b values for different models. As this thesis suggested that D^* and K have the potential to be clinically useful in subtype identification, then future work should focus on optimising the b values for IVIM and kurtosis. For the IVIM model a greater range of lower b values should be used to capture the faster diffusion and perfusion effects, for example, including 10 b values below 200mm^2/s may help capture these effects. Conversely for the kurtosis model higher b values may be needed, for example increasing the maximum b value from 1,000mm^2/s to 2,000mm^2/s.

This work could be evaluated using phantoms by only altering the b values and keeping the other parameters consistent. By repeating these experiments and selecting different b value ranges, optimisation may be possible. Following this reproducibility studies could then be performed in healthy renal tissue. Firstly, repeating the scans in a single session with no sequence alterations and then moving on to studies where field strength is altered to see if this would affect the fitted parameters.

Once the reproducibility of D^* and K are determined then Wilms' tumour subtype identification can be further explored with a larger cohort of patients. Ideally, histological slices would be matched to MR slices and on each slice regions of different cell types would be identified. Then D^* and K values could be compared between the subpopulations of distinct cellular environments, and ANOVAs and post-hoc testing could confirm significant differences.

References: Chapter 10

1. Vujanić, G. M. *et al.* Revised International Society of Paediatric Oncology (SIOP) working classification of renal tumors of childhood. *Med. Pediatr. Oncol.* **38**, 79–82 (2002).
2. Humphries, P. D., Sebire, N. J., Siegel, M. J. & Olsen, Ø. E. Tumors in Pediatric Patients at Diffusion-weighted MR Imaging: Apparent Diffusion Coefficient and Tumor Cellularity. *Radiology* **245**, 848–854 (2007).
3. Gawande, R. S., Gonzalez, G., Messing, S., Khurana, A. & Daldrup-Link, H. E. Role of diffusion-weighted imaging in differentiating benign and malignant pediatric abdominal tumors. *Pediatr. Radiol.* **43**, 836–845 (2013).
4. Aslan, M. *et al.* Diffusion-weighted MRI for differentiating Wilms tumor from neuroblastoma. *Diagn. Interv. Radiol.* **23**, 403–406 (2017).
5. McDonald, K., Sebire, N. J., Anderson, J. & Olsen, O. E. Patterns of shift in ADC distributions in abdominal tumours during chemotherapy-feasibility study. *Pediatr. Radiol.* **41**, 99–106 (2011).
6. Littooij, A. S., Humphries, P. D. & Olsen, Ø. E. Intra- and interobserver variability of whole-tumour apparent diffusion coefficient measurements in nephroblastoma: a pilot study. *Pediatr. Radiol.* **45**, 1651–1660 (2015).
7. Hales, P. W., Olsen, Ø. E., Sebire, N. J., Pritchard-Jones, K. & Clark, C. A. A multi-Gaussian model for apparent diffusion coefficient histogram analysis of Wilms' tumour subtype and response to chemotherapy. *NMR Biomed.* **28**, 948–957 (2015).
8. Littooij, A. S., Sebire, N. J. & Olsen, Ø. E. Whole-tumor apparent diffusion coefficient measurements in nephroblastoma: Can it identify blastemal predominance? *J. Magn. Reson. Imaging* **45**, 1316–1324 (2017).
9. Littooij, A. S. *et al.* Apparent diffusion coefficient as it relates to histopathology findings in post-chemotherapy nephroblastoma: a feasibility study. *Pediatr. Radiol.* **47**, 1608–1614 (2017).
10. Le Bihan, D. *et al.* Separation of diffusion and perfusion in intravoxel incoherent motion MR imaging. *Radiology* **168**, 497–505 (1988).
11. Bennett, K. M. *et al.* Characterization of continuously distributed cortical water diffusion rates with a stretched-exponential model. *Magn. Reson. Med.* **50**, 727–734 (2003).

12. Jensen, J. H., Helpern, J. A., Ramani, A., Lu, H. & Kaczynski, K. Diffusional kurtosis imaging: the quantification of non-gaussian water diffusion by means of magnetic resonance imaging. *Magn. Reson. Med.* **53**, 1432–1440 (2005).
13. Boccon-Gibod, L. *et al.* Complete necrosis induced by preoperative chemotherapy in Wilms tumor as an indicator of low risk: report of the international society of paediatric oncology (SIOP) nephroblastoma trial and study 9. *Med. Pediatr. Oncol.* **34**, 183–190 (2000).
14. Perazella, M. A. Gadolinium-contrast toxicity in patients with kidney disease: nephrotoxicity and nephrogenic systemic fibrosis. *Curr. Drug Saf.* **3**, 67–75 (2008).
15. Godzinski, J. The current status of treatment of Wilms' tumor as per the SIOP trials. *J. Indian Assoc. Pediatr. Surg.* **20**, 16–20 (2015).